

Parameter Estimation for Exponential Signals in Colored Noise Using the Pseudo-Autoregressive (PAR) Model

by

Tong-Zhang Kou

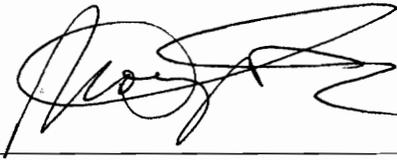
Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

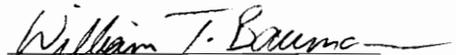
APPROVED:



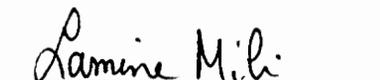
A. A. (Louis) Beex, Chairman



H. F. VanLandingham



W. T. Baumann



L. M. Mili



W. E. Kohler

April 1995
Blacksburg, Virginia

LD

5655

V856

1975

K73

C12

Parameter Estimation for Exponential Signals in Colored Noise Using the Pseudo-Autoregressive (PAR) Model

by

Tong-Zhang Kou

A. A. (Louis) Beex, Chairman

Department of Electrical Engineering

ABSTRACT

Most modern techniques for high resolution processing of closely spaced signals assume either uncorrelated noise or require knowledge of the noise covariance matrix. These assumptions are often invalid in practice. Here we propose a Pseudo-Autoregressive (PAR(M, p)) model for estimation of an arbitrary number of signals M in the presence of a p – th order autoregressive (AR) noise environment. We derive the Cramer-Rao Lower Bound (CRLB) for the parameters of damped exponential signals in the colored noise case. A closed-form expression for the Cramer-Rao Lower Bound for the Pseudo-Autoregressive (PAR(M, p)) model is obtained. Some special cases are investigated, for example, the PAR(M, p) model for $p = 0$, i.e., the white noise case, where our results agree with previous research results. We then evaluate the Cramer-Rao Lower Bound for two possibly closely spaced signals in a colored noise environment,

showing that the colored noise assumption can lead to a much lower variance bound for the exponential parameters than under the white noise assumption.

An algorithmic procedure is presented for the identification of the parameters of exponential signals, measured in colored noise. Previous papers on identifying sinusoids in noise have concentrated mainly on white noise disturbances. In a practical environment however, the disturbance is usually colored; sea-clutter in a radar context, is a lowpass type noise for example. When least squares type estimates are used in the colored noise environment, this usually leads to an unacceptable bias in the estimates. We propose an identification method, named Singular Value Decomposition Bias Elimination (SVDBE), in which it is assumed that the noise can be represented well by an AR process. The parameters of this noise model are then iteratively estimated along with the exponential signal parameters, via Singular Value Decomposition (SVD) based least squares. The iteration process starts with the white noise assumption, and improves on that by allowing the parameters in the noise model to vary away from the white noise case. A high order modal decomposition is found, and the best subset of the identified modes is selected. Simulations assess the merits of the introduced SVDBE algorithm, by comparison of the estimation results with the derived Cramer-Rao Lower Bound.

Acknowledgments

I would like to especially thank Dr. A. A. (Louis) Beex for his patience and readiness when dealing with me for the last few years in discussing the questions I had when doing the research for this dissertation. The guidance and the technical insight provided by Dr. Beex have earned my respect and admiration. His suggestions were always helpful. I would like to thank Dr. VanLandingham, Dr. Baumann, Dr. Mili and Dr. Kohler for their advice and their judgments as my dissertation committee.

A special thanks to my parents, my wife Huo Yu Ying, my son Kou Lei and my brothers and sisters. I would like to thank them for their support and encouragement during the years I spent in graduate school.

Finally I would like to thank the Continuing Education Center at Virginia Tech, and thank a lot of my friends from Virginia Tech.

Table of Contents

Chapter 1	Introduction	1
Chapter 2	Pseudo-Autoregressive Model (PAR)	7
Chapter 3	The Cramer-Rao Lower Bound for the Pseudo-Autoregressive (PAR) Model	15
Chapter 4	Special Case CRLB for the PAR Model	33
4.1	Multiple Experiments in White Noise	33
4.2	Single Experiment in White Noise	35
4.3	Single Experiment in Colored Noise	37
4.3.1	Colored Noise of Order 0 (White Noise)	42
4.3.2	The First Order Colored Noise Case	42
Chapter 5	Simulated CRLB for PAR Model	45
5.1	First Order Autoregressive Noise With a Single Signal	46
5.1.1	CRLB (\hat{f}) for AR(1) Noise for Different Data Length N	46

5.1.2	Different Bandwidths for Same Power AR(1) Noise	51
5.2	First Order Autoregressive Noise With Two Signals	54
Chapter 6	Analytical CRLB for PAR Model	61
6.1	PAR Model Structure Analysis	61
6.2	Analytical CRLB for Autoregressive Colored Noise Case	68
6.3	Analytical CRLB Evaluation Results	70
6.3.1	First Order Autoregressive Noise With a Single Signal	70
6.3.2	First Order Autoregressive Noise With Two Signals	74
Chapter 7	Parameter Estimation of Exponential Signals in Colored Noise	78
7.1	Introduction of the Estimation Method	79
7.2	Bias Due to Colored Noise	82
7.3	The SVDBE Algorithm	84
7.4	Mode Subset Selection	90
7.4.1	Mode Subset Selection in the White Noise Case	91

7.4.2	Best Subset Selection in the Colored Noise Case	92
7.5	Simulation Results	95
7.5.1	Signal Frequency Estimation in the White Noise Case	95
7.5.2	Best Subset Selection of Modes in the White Noise Case	97
7.5.3	Signal Frequency Estimation in the AR Noise Case	99
7.5.4	Effect of Signal Frequency Difference in the AR Noise Case	102
7.5.5	Mode Subset Selection in the AR Noise Case	104
Chapter 8	Conclusions	107
Appendix A		110
Appendix B		114
Appendix C		124
Appendix D		150
Appendix E		161
References		168
Vita		177

LIST OF FIGURES

Fig. 5.1a	Power spectral densities for AR(1) noise and white noise of same power.	47
Fig. 5.1b	CRLB comparison for AR(1) noise and white noise for $N=5$.	48
Fig. 5.1c	CRLB comparison for AR(1) noise and white noise for $N=5$; mean value based on 50×50 realizations.	48
Fig. 5.1d	CRLB comparison for AR(1) noise and white noise for $N=10$.	49
Fig. 5.1e	CRLB comparison for AR(1) noise and white noise for $N=10$; mean value based on 50×50 realizations.	49
Fig. 5.1f	CRLB comparison for AR(1) noise and white noise for $N=25$.	50
Fig. 5.1g	CRLB comparison for AR(1) noise and white noise for $N=25$; mean value based on 50×50 realizations.	50
Fig. 5.2a	Spectrum of AR noise with different d_1 and white noise spectrum, same power in all.	52
Fig. 5.2b	$\text{CRLB}(\hat{f})$ for AR noise with different d_1 and $\text{CRLB}(\hat{f})$ for white noise, mean value based on 50×50 realizations.	53
Fig. 5.3	Locations of signal frequencies and colored noise.	55

Fig. 5.4	CRLB(\hat{f}) versus signal frequency difference Δf for AR(1) noise.	56
Fig. 5.5a	CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for N=10 and 50×50 realizations.	59
Fig. 5.5b	CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for N=10, mean value from 50×50 realizations.	59
Fig. 5.6a	CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for N=25 and 50×50 realizations.	60
Fig. 5.6b	CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for N=25, mean value from 50×50 realizations.	60
Fig. 6.1	Pseudo-Autoregressive (PAR) model structure.	62
Fig. 6.2	Simplified Pseudo-Autoregressive (PAR) model structure.	64
Fig. 6.3	Analytical CRLB for AR(1) noise and white noise for N=10.	72
Fig. 6.4	Analytical CRLB for AR(1) noise and white noise for N=25.	73
Fig. 6.5	Analytical CRLB for AR(1) noise and white noise vs. signal frequency difference Δf for N=10.	76
Fig. 6.6	Analytical CRLB for AR(1) noise and white noise vs. signal frequency difference Δf for N=25.	77
Fig. 7.1	Minimum prediction error versus subset size for white noise.	98

Fig. 7.2	Root mean square error for frequency f_1 estimation in AR(1) noise versus SVDBE iteration number.	101
Fig. 7.3	Variance for frequency f_1 estimation and CRLB for frequency f_1 estimation for AR(1) noise and CRLB for white noise versus signal frequency Δf for $N = 25$.	103
Fig. 7.4	The best subset selection for the autoregressive colored noise case.	105
Fig. 7.5	The best subset selection for the incorrect white noise assumption.	106

LIST OF TABLES

Table 1 Root mean square error of frequency f_1 estimation
 versus signal-to-noise ratio.

96

1. Introduction

The problem of estimating model parameters from noisy measurement data is an important active area of research in the field of signal processing. Recently, high resolution methods for such estimation have received the attention of many researchers [1-12, 31-40]. This problem is of so much interest because it has a broad range of useful applications, such as radar, sonar, communications, and seismic exploration. Among these applications are direction finding with narrow-band sensor arrays [3], moving target detection and tracking [25, 26, 27], passive sonar array processing [28, 29], underwater acoustics data processing [30], resolution of overlapping echoes [31], and array signal processing in seismology and geophysics [32]. These problems can be reduced to estimating parameters of random signals in additive noise. The physical phenomena are often characterized by stochastic processes; the simplest such process is a white process for which the samples at different times are uncorrelated. If a stochastic process is correlated in time, modeling of the correlation structure is often useful in analyzing the process. High resolution methods usually assume that the background noise is spatially

uncorrelated or that the covariance matrix of the noise is known, i.e., the white noise case. In practice, however, these assumptions are often not valid. This has led to the consideration of estimation methods based on a correlated noise field, i.e., the colored noise environment. In searching for the best estimates it is a good idea to be able to bound the performance of an estimator. In practice we are usually interested in unbiased estimators. The variance of any unbiased estimators is bounded by the Cramer-Rao Lower Bound [13,14].

Estimation of random signals in additive noise is encountered in diverse areas of engineering. The nature of the signals or the complexity of the propagation media are the main reason for causing randomness in the received signals. Usually the received data consists of a weak signal immersed in a strong background noise, thereby presenting a difficult scenario for estimation. An important problem in signal processing is to combat the noise, so that the signal information can be enhanced, or extracted.

A number of algorithms exist for estimating unknown signal parameters from measured output of a sensor array. A lot of the recent work in array processing has focused on methods for high-resolution direction-of-arrival estimation. When the emitter signals are generated by spatially close sources, conventional beam forming methods fail to give satisfactory answers. A number of different methods have been proposed for estimating closely spaced signals. Many of these techniques can be formulated in a subspace fitting structure. From this viewpoint the connections between different

algorithms become clear. Subspace or eigenvector methods are known to have high resolution capabilities and yield accurate estimates in the white noise case. The multiple signal classification (MUSIC) algorithm [3, 33, 34] with a one-dimensional search, provides a nice geometric interpretation of the direction-of-arrival problem and has received much attention. The asymptotic properties of MUSIC have been shown in the literature [2, 15, 35].

The pencil method [38] is original to estimation of signal parameters via the rotation invariant technique (ESPRIT) [6, 36]. In the single channel case, ESPRIT is based on the observation that in a uniformly sampled time or space series there is a constant delay and attenuation between any two adjacent samples. We can define two data matrices, one as the original, another as a delayed data matrix. On the basis of these two data matrices, a pencil of matrices is formed, and the roots of the pencil of matrices can be used to estimate the unknown parameters. The total least squares variation of ESPRIT can be found in the literature [37]. The pencil method and ESPRIT belong to the class of signal subspace methods in which one relies on an eigendecomposition of the sample covariance matrix. These methods have their origin in Pisarenko's method [39] for harmonic retrieval.

Maximum likelihood theory has been applied to a number of fundamental problems in signal processing. It is now an active area of research for time series identification, spectrum analysis, and sensor array processing. Many authors have applied maximum

likelihood theory or its equivalent, least-squares theory, to estimate the parameters for deterministic signals observed in additive Gaussian noise. Examples of maximum likelihood theory applied to time series analysis can be found in the literature [40, 42, 43], including an algorithm for pole-zero modeling and analysis [4], and the use of exact maximum likelihood estimation for superimposed exponential signals in white noise [5].

The early work on estimation of structured covariance matrices [41] led to using maximum likelihood estimation for Toeplitz matrices and the use of such estimators in autoregressive spectral estimation. The properties of maximum likelihood estimates of structured covariance matrices have been studied extensively [44, 45].

The methods discussed above are all closely related, although the actual computations involved can be quite different. The above mentioned estimates assume the noise spectrum to be either white or known up to a scaling factor. Unfortunately, this is an unrealistic assumption in many practical situations. Recently the colored background noise cases have received more attention [19, 47, 48, 49, 50]. Some approaches are suboptimal [49], others require additional assumptions [48]. Some examine the real valued data case for frequency estimation in colored noise [50], others require a very high computational burden [46].

If a noise model can be estimated, then the colored noise problem can be converted to the standard white noise problem by a whitening approach. Recent simulations, comparing the results obtained from the MUSIC method for source bearing estimation,

based on data with and without whitening [46] have shown the whitening approach to be effective when a large number of snapshots is available. The whitening approach failed to converge in only one out of 10 trials in which 300 snapshots were used [46]. The failure rate is expected to increase significantly when much shorter records are used, such as on the order of 25 snapshots. It is exactly the latter situation for which we aim to improve the estimation process, not by whitening, by explicitly incorporating the estimation of the noise model into the estimation process.

The Cramer-Rao Lower Bound provides an effective measure for evaluating estimates of model parameters. The bounds computed in the literature [15, 16, 17, 18, 24] use the white noise assumption. No comprehensive study of performance bounds for damped exponential signals in colored noise has been published.

The contributions of this dissertation relate to the estimation of exponential signal frequencies from short records, in a colored noise environment. The Pseudo-Autoregressive (PAR(M, p)) model for estimation of an arbitrary number, M , of exponential signals in the presence of a p -th order autoregressive (AR) noise environment is proposed. The Cramer-Rao Lower Bound (CRLB) for parameter estimation under the PAR model is derived. A significant finding is that the Cramer-Rao Lower Bound for the colored noise environment can be much lower than that under the white noise assumption, away from the noise peak. We also propose an identification method, named Singular Value Decomposition Bias Elimination (SVDBE), which assumes

that the noise can be represented well by an AR process. The parameters of this noise model are then iteratively estimated along with the exponential signal parameters, via Singular Value Decomposition (SVD) based least squares. The iteration process starts with the white noise assumption, and improves on that by allowing the parameters in the noise model to vary away from the white noise case. A high order modal decomposition is found, and a new criterion proposed for the successful determination of the best subset of the identified modes in the colored noise case. Simulations show the merits of the SVDBE algorithm, in that the estimation results are close to the derived Cramer-Rao Lower Bound.

The dissertation is organized as follows. Chapter 2 introduces the Pseudo-Autoregressive (PAR) model for the sum of damped exponential signals in a colored noise environment. Chapter 3 analyzes the Cramer-Rao Lower Bound for the Pseudo-Autoregressive model. In Chapter 4 we discuss some special cases of the derived Cramer-Rao Lower Bound. Chapter 5 gives simulated results for the CRLB. In Chapter 6 the analytical Cramer-Rao Lower Bound for the PAR model is obtained and compared with the simulated case introduced in Chapter 3. In Chapter 7 the Singular Value Decomposition Bias Elimination (SVDBE) algorithm is proposed for estimation of signal parameters in AR(1) colored noise. The problem of the best subset size selection is discussed here too. The conclusions are given in Chapter 8.

2. Pseudo-Autoregressive Model (PAR)

The observations are a short record of the data sequence y_n , $n = 1, 2, 3, \dots, N$, which is assumed to be composed of uniformly spaced samples of a sum of exponential signals, s_n , and a colored measurement noise v_n .

$$y_n = s_n + v_n \quad n = 1, 2, 3, \dots, N \quad (2-1a)$$

where

$$s_n = \sum_{k=1}^M a_k (\rho_k)^n \exp(j\phi_k) \quad (2-1b)$$

$$= \sum_{k=1}^M \bar{a}_k \rho_k^n \quad (2-1c)$$

The complex amplitude \bar{a}_k is given by

$$\bar{a}_k = a_k \exp(j\phi_k) \quad (2-1d)$$

and the complex mode ρ_k is given by

$$\rho_k = \exp(-\alpha_k + j\omega_k) \quad (2-1e)$$

where a_k is the amplitude, α_k is the damping factor, ω_k is the radian frequency, and ϕ_k is the phase of the k – th exponential signal term. In general the value of M is unknown. However, let us first assume that the value of M is known, and assume the maximum possible value of M is equal to half the number of data points, i. e., $M = N / 2$. Of particular interest and difficulty is the situation where the observation record is short and the signal frequency separations are less than $1 / N$, i.e., the difference in the signal frequencies is less than the Rayleigh limit.

Assume that the colored noise v_n has autoregressive structure, and therefore satisfies

$$v_n = -\sum_{i=1}^p d_i v_{n-i} + e_n \quad (2-2)$$

where p is the order of the AR process, e_n is an independent identically distributed complex Gaussian sequence with zero mean and variance σ_e^2 , and $D(z)$ is minimum phase.

$D(z)$ is defined as

$$D(z) = \sum_{i=0}^p d_i z^{-i} \quad , \quad d_0 \equiv 1 \quad (2-3)$$

Now (2-1a) can be written as

$$\begin{aligned} y_n &= s_n - \sum_{i=1}^p d_i v_{n-i} + e_n \\ &= s_n - \sum_{i=1}^p d_i (y_{n-i} - s_{n-i}) + e_n \\ &= s_n + \sum_{i=1}^p d_i s_{n-i} - \sum_{i=1}^p d_i y_{n-i} + e_n \\ &= \tilde{s}_n - \sum_{i=1}^p d_i y_{n-i} + e_n \end{aligned} \quad (2-4)$$

Here the signal \tilde{s}_n is the signal s_n filtered by the MA filter with system function

$$H_{MA}(z) = D(z) \quad (2-5)$$

Let us see what happens to the signal after the MA filter operates on it. From (2-1b)

$$\begin{aligned} s_n &= \sum_{k=1}^M a_k (\rho_k)^n \exp(j\phi_k) \\ &= \sum_{k=1}^M \bar{a}_k \exp((- \alpha_k + j\omega_k)n) \end{aligned} \quad (2-6)$$

After the MA filter $H_{MA}(z)$ operates on the signal, we get a filtered signal \tilde{s}_n as follows

$$\begin{aligned} \tilde{s}_n &= \sum_{k=1}^M \bar{a}_k D(z) \Big|_{z=\exp(-\alpha_k + j\omega_k)} \exp((- \alpha_k + j\omega_k)n) \\ &= \sum_{k=1}^M \hat{a}_k \exp((- \alpha_k + j\omega_k)n) \end{aligned} \quad (2-7)$$

where we have defined \hat{a}_k as follows

$$\hat{a}_k = \bar{a}_k D(z) \Big|_{z=\exp(-\alpha_k + j\omega_k)} \quad (2-8)$$

We see that the damping factors in \tilde{s}_n are still α_k , and the frequencies are still ω_k , for the k -th signal component of \tilde{s}_n . Substitution of (2-7) into (2-4) yields

$$y_n = \sum_{k=1}^M \hat{a}_k \exp[-\alpha_k n + j\omega_k n] - \sum_{i=1}^p d_i y_{n-i} + e_n \quad (2-9)$$

We propose the structure in (2-9) as the Pseudo-Autoregressive (PAR) model for the measurement process y_n . The PAR(M, p) model has unknown parameters

$$\xi = M, \{\hat{a}_k, \alpha_k, \omega_k\}_1^M, p, \{d_i\}_1^p, \sigma_e^2 \quad (2-10)$$

Note that for $M = 0$ the PAR($0, p$) model reduces to a p -th order autoregressive model. Note that when v_n in (2-1a) is white noise, then in the above PAR(M, p) model, all $d_i \equiv 0$, and $\hat{a}_k \equiv \bar{a}_k$, while the damping factors and frequencies remain the original ones for any AR(p) noise.

For the exponential signal in (2-1) and the noise v_n an actual AR(p) process, the observations y_n obey the PAR(M, p) model. If the noise does not have a pure AR structure, then the observation sequence y_n does not satisfy the PAR model exactly for any finite order p , even if M is correct. The PAR model is expected to produce better

signal parameter estimates for the case of colored noise than with the assumption that the noise is white. We will illustrate this point with simulation examples in Chapter 4.

In Appendix A the conditional probability density function of the Pseudo-Autoregressive observation sequence is derived, as

$$P(y_{p+1}, y_{p+2}, \dots, y_N | y_1, y_2, \dots, y_p; \xi)$$

$$= \left(\frac{1}{\pi \sigma_e^2} \right)^{(N-p)} \exp \left\{ -\frac{1}{\sigma_e^2} \sum_{n=p+1}^N \left| y_n - \sum_{k=1}^M \hat{a}_k \exp(-\alpha_k n + j\omega_k n) + \sum_{i=1}^p d_i y_{n-i} \right|^2 \right\} \quad (2-11)$$

Now let us derive the matrix form of the PAR(M, p) model in (2-9), for $n = p+1, p+2, \dots, N$.

For $n = p+1$ in (2-9) we get

$$\begin{aligned} y_{p+1} &= \sum_{k=1}^M e^{(-\alpha_k + j\omega_k)p} \hat{a}_k e^{(-\alpha_k + j\omega_k)} - \sum_{i=1}^p d_i y_{p+1-i} + e_{p+1} \\ &= \sum_{k=1}^M \rho_k^p \tilde{a}_k - \sum_{i=1}^p d_i y_{p+1-i} + e_{p+1} \end{aligned}$$

$$= \begin{bmatrix} \rho_1^p & \rho_2^p & \dots & \rho_M^p \end{bmatrix} \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \\ \vdots \\ \tilde{a}_M \end{bmatrix} - \begin{bmatrix} y_p & y_{p-1} & \dots & y_1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix} + e_{p+1} \quad (2-12)$$

where we defined

$$\tilde{a}_i = \hat{a}_i e^{(-\alpha_i + j\omega_i)} \quad (2-13)$$

In a similar rewrite we get, for $n = p + 2$,

$$\begin{aligned} y_{p+2} &= \sum_{k=1}^M e^{(-\alpha_k + j\omega_k)(p+1)} \hat{a}_k e^{(-\alpha_k + j\omega_k)} - \sum_{i=1}^p d_i y_{p+2-i} + e_{p+2} \\ &= \sum_{k=1}^M \rho_k^{p+1} \tilde{a}_k - \sum_{i=1}^p d_i y_{p+2-i} + e_{p+2} \\ &= \begin{bmatrix} \rho_1^{p+1} & \rho_2^{p+1} & \dots & \rho_M^{p+1} \end{bmatrix} \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \\ \vdots \\ \tilde{a}_M \end{bmatrix} - \begin{bmatrix} y_{p+1} & y_p & \dots & y_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix} + e_{p+2} \end{aligned} \quad (2-14)$$

Collecting all the above for $n = p + 1, p + 2, \dots, N$, yields

$$\begin{bmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \rho_1^p & \rho_2^p & \dots & \rho_M^p \\ \rho_1^{p+1} & \rho_2^{p+1} & \dots & \rho_M^p \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1^{N-1} & \rho_2^{N-1} & \dots & \rho_M^{N-1} \end{bmatrix} \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \\ \vdots \\ \tilde{a}_M \end{bmatrix} - \begin{bmatrix} y_p & y_{p-1} & \dots & y_1 \\ y_{p+1} & y_p & \dots & y_2 \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-1} & y_{N-2} & \dots & y_{N-p} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix} + \begin{bmatrix} e_{p+1} \\ e_{p+2} \\ \vdots \\ e_N \end{bmatrix} \quad (2-15a)$$

$$\mathbf{y} = \mathbf{R} \tilde{\mathbf{a}} - \mathbf{Y} \mathbf{d} + \mathbf{e} \quad (2-15b)$$

In matrix form we thus obtain

$$\mathbf{y} = \mathbf{R}\tilde{\mathbf{a}} - \mathbf{Y}\mathbf{d} + \mathbf{e} \quad (2-16)$$

with definitions obvious from the equivalence of (2-15a) and (2-15b).

Equation (2-15) is the matrix form of the PAR(M, p) model. Based on this PAR(M, p) model we can analyze the accuracy of parameter estimation. The Cramer-Rao Lower Bound for estimating unknown parameters is investigated in the next chapter.

3. The Cramer-Rao Lower Bound for the Pseudo-Autoregressive (PAR) model

To evaluate the accuracy of parameter estimates, a general expression for the Cramer-Rao Lower Bound (CRLB) is now derived. For the white noise case, i.e. v_n is white, the Cramer-Rao Lower Bound can be found in several places [15, 16]. In this chapter we will develop the Cramer-Rao Lower Bound for the colored noise environment. We will derive the Cramer-Rao Lower Bound for PAR(M, p) model parameter estimation. In the estimation process the Cramer-Rao Lower Bound serves as an indicator of the best unbiased estimation that can be made from the available observations. It is important, therefore, to have knowledge of the Cramer-Rao Lower Bound, in order to assess the performance of different estimators.

From the probability density function of the measurement process PAR(M, p) model in (2-11), and (2-16), we get the likelihood function $L(\cdot)$ for Q experiments, each consisting of $(N - p)$ observations.

$$\begin{aligned}
L(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(Q)) &= \frac{1}{(\pi)^{(N-p)Q} (\sigma_e^2)^{(N-p)Q}} \exp\left(-\frac{1}{\sigma_e^2} \sum_{i=1}^Q (\mathbf{e}^H(i)\mathbf{e}(i))\right) \\
&= \frac{1}{(\pi)^{(N-p)Q} (\sigma_e^2)^{(N-p)Q}} \exp\left\{-\frac{1}{\sigma_e^2} \sum_{i=1}^Q \left[(\mathbf{y}(i) - R\tilde{\mathbf{a}}(i) + Y(i)\mathbf{d})^H (\mathbf{y}(i) - R\tilde{\mathbf{a}}(i) + Y(i)\mathbf{d}) \right]\right\}
\end{aligned} \tag{3-1}$$

The case $Q=1$ corresponds to the single-experiment problem with uniform sampling. The case $Q > 1$ corresponds to a multiple experiment with a time series. For example, in the multiple experiment case the components of the measurement vector ("snapshot") represent the output from N individual sensors for a linear uniform narrow-band array processing problem. From snapshot to snapshot the frequencies ω_k and damping coefficients α_k do not vary, but the amplitudes α_k may. This type of data is found in the sensor array problem [5, 6, 15, 16, 23].

The log likelihood function, therefore, is

$$\begin{aligned}
\ln L &= -(N-p)Q \ln(\pi) - (N-p)Q \ln(\sigma_e^2) \\
&\quad - \frac{1}{\sigma_e^2} \sum_{i=1}^Q \left[(\mathbf{y}(i) - R\tilde{\mathbf{a}}(i) + Y(i)\mathbf{d})^H (\mathbf{y}(i) - R\tilde{\mathbf{a}}(i) + Y(i)\mathbf{d}) \right]
\end{aligned} \tag{3-2}$$

Defining

$$\tilde{\mathbf{a}}_R(i) = \text{Re}(\tilde{\mathbf{a}}(i)) \quad (3-3a)$$

$$\tilde{\mathbf{a}}_I(i) = \text{Im}(\tilde{\mathbf{a}}(i)) \quad (3-3b)$$

we now evaluate the derivatives of the log likelihood function with respect to the variables of interest: $\sigma_e^2, \tilde{\mathbf{a}}_R(i), \tilde{\mathbf{a}}_I(i), \mathbf{d}_R, \mathbf{d}_I, \alpha$ and ω . We obtain directly from (3-2)

$$\frac{\partial \ln L}{\partial \sigma_e^2} = -\frac{(N-p)Q}{\sigma_e^2} + \frac{1}{\sigma_e^2} \sum_{i=1}^Q \mathbf{e}^H(i) \mathbf{e}(i) \quad (3-4)$$

Appendix B gives the derivation of the following results:

$$\frac{\partial \ln L}{\partial \mathbf{d}_R} = -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \text{Re}[Y^H(k) \mathbf{e}(k)] \quad (3-5)$$

$$\frac{\partial \ln L}{\partial \mathbf{d}_I} = -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \text{Im}[Y^H(k) \mathbf{e}(k)] \quad (3-6)$$

$$\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} = \frac{2}{\sigma_e^2} \text{Re} \left[R^H \mathbf{e}(k) \right] \quad k = 1, 2, \dots, Q \quad (3-7)$$

$$\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)} = \frac{2}{\sigma_e^2} \text{Im} \left[R^H \mathbf{e}(k) \right] \quad k = 1, 2, \dots, Q \quad (3-8)$$

$$\frac{\partial \ln L}{\partial \alpha_l} = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \text{Re} \left\{ \tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \alpha_l} \mathbf{e}(i) + \frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha_l} R^H \mathbf{e}(i) \right\} \quad l = 1, 2, \dots, M \quad (3-9)$$

The matrix formulation of (3-9) is

$$\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \text{Re} \left\{ A_{diag}^H(i) R_{\boldsymbol{\alpha}}^H \mathbf{e}(i) + A_{diag, \boldsymbol{\alpha}}^*(i) R^H \mathbf{e}(i) \right\} \quad (3-10)$$

where we defined the damping parameter vector $\boldsymbol{\alpha}$

$$\boldsymbol{\alpha} = [\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_M]^T \quad (3-11)$$

$$A_{diag}(i) = \text{diag}[\tilde{a}_1(i) \quad \tilde{a}_2(i) \quad \dots \quad \tilde{a}_M(i)] \quad i = 1, 2, \dots, Q \quad (3-12)$$

$$R_{\boldsymbol{\alpha}} = \begin{bmatrix} \frac{\partial \rho_1}{\partial \alpha_1} & \frac{\partial \rho_2}{\partial \alpha_2} & \dots & \frac{\partial \rho_M}{\partial \alpha_M} \end{bmatrix} \quad (3-13)$$

$$\boldsymbol{\rho}_l = [\rho_l^p \quad \rho_l^{p+1} \quad \cdots \quad \rho_l^{N-1}]^T \quad l = 1, 2, \dots, M \quad (3-14)$$

$$R = [\boldsymbol{\rho}_1 \quad \boldsymbol{\rho}_2 \quad \cdots \quad \boldsymbol{\rho}_M] \quad (3-15)$$

Note that $A_{diag,\alpha}(i)$ is a column derivative with respect to the elements of α , analogous to R_α in (3-13).

For the frequency parameter derivatives we find similarly

$$\frac{\partial \ln L}{\partial \omega_l} = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \operatorname{Re} \left[\tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \omega_l} \mathbf{e}(i) + \frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega_l} R^H \mathbf{e}(i) \right] \quad l=1,2,\dots,M \quad (3-16)$$

which can be rewritten in matrix form as

$$\frac{\partial \ln L}{\partial \boldsymbol{\omega}} = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \operatorname{Re} \left[A_{diag}^H(i) R_\omega^H \mathbf{e}(i) + A_{diag,\omega}^*(i) R^H \mathbf{e}(i) \right] \quad (3-17)$$

where

$$\boldsymbol{\omega} = [\omega_1 \quad \omega_2 \quad \cdots \quad \omega_M]^T \quad (3-18)$$

$$R_{\omega} = \begin{bmatrix} \frac{\partial \rho_1}{\partial \omega_1} & \frac{\partial \rho_2}{\partial \omega_2} & \dots & \frac{\partial \rho_M}{\partial \omega_M} \end{bmatrix} \quad (3-19)$$

From these results, the Fisher information matrix can be evaluated as

$$I_f = E[\psi\psi^T] \quad (3-20a)$$

where

$$\psi^T = \frac{\partial \ln L}{\partial \left[\sigma_s^2 \quad \tilde{\mathbf{a}}_R^T(1) \quad \tilde{\mathbf{a}}_I^T(1) \quad \dots \quad \tilde{\mathbf{a}}_R^T(Q) \quad \tilde{\mathbf{a}}_I^T(Q) \quad \mathbf{d}_R^T \quad \mathbf{d}_I^T \quad \boldsymbol{\alpha}^T \quad \boldsymbol{\omega}^T \right]} \quad (3-20b)$$

and $E[\cdot]$ denotes expectation. We now need to find the elements of I_f , such as

$$E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}} \right]^T \right\} = -E \left\{ \frac{\partial}{\partial \tilde{\mathbf{a}}} \left(\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}} \right)^T \right\} \quad (3-21)$$

which for later convenience will be given compact letter definitions as well.

In order to evaluate these expectations we need some results which were proven earlier [15]:

$$E[\mathbf{e}^H(i)\mathbf{e}(i)\mathbf{e}^H(l)\mathbf{e}(l)] = \begin{cases} (N-p)^2\sigma_e^4 & i \neq l \\ (N-p)(N-p+1)\sigma_e^4 & i = l \end{cases} \quad (3-22)$$

$$E[\mathbf{e}^H(i)\mathbf{e}(i)\mathbf{e}^T(l)] = \mathbf{0}^T \quad \text{for all } i \text{ and } l \quad (3-23)$$

$$\operatorname{Re}\{x\}\operatorname{Re}\{y^T\} = \frac{1}{2}\{\operatorname{Re}[xy^T] + \operatorname{Re}[xy^H]\}$$

$$\operatorname{Im}\{x\}\operatorname{Im}\{y^T\} = -\frac{1}{2}\{\operatorname{Re}[xy^T] - \operatorname{Re}[xy^H]\}$$

$$\operatorname{Re}\{x\}\operatorname{Im}\{y^T\} = \frac{1}{2}\{\operatorname{Im}[xy^T] - \operatorname{Im}[xy^H]\}$$

We will also make use of the following:

$$\operatorname{Im}\{x\}\operatorname{Re}\{y^T\} = \frac{1}{2}\{\operatorname{Im}[xy^T] + \operatorname{Im}[xy^H]\} \quad (3-24)$$

By using equation (3-22), we obtain

$$\begin{aligned}
E\left[\frac{\partial \ln L}{\partial \sigma_e^2}\right]^2 &= \\
&= \frac{Q^2(N-p)^2}{\sigma_e^4} - 2\frac{Q(N-p)}{\sigma_e^6} \sum_{i=1}^Q E[\mathbf{e}(i)^H \mathbf{e}(i)] + \frac{1}{\sigma_e^8} \sum_{i=1}^Q \sum_{l=1}^Q E[\mathbf{e}^H(i)\mathbf{e}(i)\mathbf{e}^H(l)\mathbf{e}(l)] \\
&= \frac{(N-p)^2 Q^2}{\sigma_e^4} - 2\frac{(N-p)^2 Q^2}{\sigma_e^4} + \frac{Q}{\sigma_e^4} [(N-p)(N-p+1) + (Q-1)(N-p)^2] \\
&= \frac{Q^2(N-p)^2}{\sigma_e^4} - 2\frac{Q^2(N-p)^2}{\sigma_e^4} + \frac{Q^2(N-p)^2}{\sigma_e^4} + \frac{(N-p)Q}{\sigma_e^4} \\
&= \frac{(N-p)Q}{\sigma_e^4} \tag{3-25}
\end{aligned}$$

From equation (3-23) and equation (3-4) we note that $\frac{\partial \ln L}{\partial \sigma_e^2}$ is not correlated

with any of the other partial derivatives.

Next, we use equations (3-21) and (3-24) and the fact that $E[\mathbf{e}(i)\mathbf{e}^T(l)] = 0$ for all

i and l and get

$$\overline{C}^{\Delta} = E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(i)} \right]^T \right\} = \frac{2}{\sigma_e^2} \text{Re} [R^H R] \delta_{k,i} \quad (3-26a)$$

$$\overline{C}^{\Delta} = E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(i)} \right]^T \right\} = \frac{2}{\sigma_e^2} \text{Im} [R^H R] \delta_{k,i} \quad (3-26b)$$

$$E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(i)} \right]^T \right\} = -\frac{2}{\sigma_e^2} \text{Im} [R^H R] \delta_{k,i} = -\overline{C} \quad (3-27a)$$

$$E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(i)} \right]^T \right\} = \frac{2}{\sigma_e^2} \text{Re} [R^H R] \delta_{k,i} = \overline{C} \quad (3-27b)$$

$$\overline{B}(k)^{\Delta} = E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \mathbf{d}_R} \right]^T \right\} = -\frac{2}{\sigma_e^2} \text{Re} [R^H E(Y(k))] \quad (3-28a)$$

$$\overline{B}(k)^{\Delta} = E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)} \right] \left[\frac{\partial \ln L}{\partial \mathbf{d}_R} \right]^T \right\} = -\frac{2}{\sigma_e^2} \text{Im} [R^H E(Y(k))] \quad (3-28b)$$

$$E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \mathbf{d}_I} \right]^T \right\} = \frac{2}{\sigma_e^2} \text{Im} [R^H E(Y(k))] = -\overline{B}(k) \quad (3-29a)$$

$$E\left\{\left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)}\right]\left[\frac{\partial \ln L}{\partial \mathbf{d}_I}\right]^T\right\} = -\frac{2}{\sigma_e^2} \operatorname{Re}\left[R^H E(Y(k))\right] = \overline{B(k)} \quad (3-29b)$$

$$\overline{F}^\Delta = E\left\{\left[\frac{\partial \ln L}{\partial \mathbf{d}_R}\right]\left[\frac{\partial \ln L}{\partial \mathbf{d}_R}\right]^T\right\} = \frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Re}\left[E(Y^H(k)Y(k))\right] \quad (3-30a)$$

$$\overline{F}^\Delta = E\left\{\left[\frac{\partial \ln L}{\partial \mathbf{d}_I}\right]\left[\frac{\partial \ln L}{\partial \mathbf{d}_R}\right]^T\right\} = \frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Im}\left[E(Y^H(k)Y(k))\right] \quad (3-30b)$$

$$E\left\{\left[\frac{\partial \ln L}{\partial \mathbf{d}_R}\right]\left[\frac{\partial \ln L}{\partial \mathbf{d}_I}\right]^T\right\} = -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Im}\left[E(Y^H(k)Y(k))\right] = -\overline{F} \quad (3-31a)$$

$$E\left\{\left[\frac{\partial \ln L}{\partial \mathbf{d}_I}\right]\left[\frac{\partial \ln L}{\partial \mathbf{d}_I}\right]^T\right\} = \frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Re}\left[E(Y^H(k)Y(k))\right] = \overline{F} \quad (3-31b)$$

$$\begin{aligned} \overline{W(k)}^\Delta &= E\left\{\left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)}\right]\left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}}\right]^T\right\} \\ &= \frac{2}{\sigma_e^2} \left\{ \operatorname{Re}\left[R^H R A_{diag,\omega}(k)\right] + \operatorname{Re}\left[R^H R_\omega A_{diag}(k)\right] \right\} \end{aligned} \quad (3-32a)$$

$$\begin{aligned} \overline{W(k)}^{\Delta} &= E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right]^T \right\} \\ &= \frac{2}{\sigma_e^2} \left\{ \text{Im} \left[R^H R A_{diag, \omega}(k) \right] + \text{Im} \left[R^H R_{\omega} A_{diag}(k) \right] \right\} \end{aligned} \quad (3-32b)$$

$$\begin{aligned} \overline{V}^{\Delta} &= E \left\{ \left[\frac{\partial \ln L}{\partial \mathbf{d}_R} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right]^T \right\} \\ &= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \left\{ \text{Re} \left[E(Y^H(k)) R A_{diag, \omega}(k) \right] + \text{Re} \left[E(Y^H(k)) R_{\omega} A_{diag}(k) \right] \right\} \end{aligned} \quad (3-33a)$$

$$\begin{aligned} \overline{V}^{\Delta} &= E \left\{ \left[\frac{\partial \ln L}{\partial \mathbf{d}_I} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right]^T \right\} \\ &= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \left\{ \text{Im} \left[E(Y^H(k)) R A_{diag, \omega}(k) \right] + \text{Im} \left[E(Y^H(k)) R_{\omega} A_{diag}(k) \right] \right\} \end{aligned} \quad (3-33b)$$

$$\overline{S(k)}^{\Delta} = E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\}$$

$$= \frac{2}{\sigma_e^2} \left\{ \operatorname{Re} \left[R^H R A_{diag,\alpha}(k) \right] + \operatorname{Re} \left[R^H R_\alpha A_{diag}(k) \right] \right\} \quad (3-34a)$$

$$\overline{S(k)}^\Delta = E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(k)} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\}$$

$$= \frac{2}{\sigma_e^2} \left\{ \operatorname{Im} \left[R^H R A_{diag,\alpha}(k) \right] + \operatorname{Im} \left[R^H R_\alpha A_{diag}(k) \right] \right\} \quad (3-34b)$$

$$\overline{\Gamma}^\Delta = E \left\{ \left[\frac{\partial \ln L}{\partial \mathbf{d}_R} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\}$$

$$= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \left\{ \operatorname{Re} \left[E(Y^H(k)) R A_{diag,\alpha}(k) \right] + \operatorname{Re} \left[E(Y^H(k)) R_\alpha A_{diag}(k) \right] \right\} \quad (3-35a)$$

$$\overline{\Gamma}^\Delta = E \left\{ \left[\frac{\partial \ln L}{\partial \mathbf{d}_I} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\}$$

$$= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \left\{ \operatorname{Im} \left[E(Y^H(k)) R A_{diag,\alpha}(k) \right] + \operatorname{Im} \left[E(Y^H(k)) R_\alpha A_{diag}(k) \right] \right\} \quad (3-35b)$$

$$\begin{aligned}
X &\stackrel{\Delta}{=} E \left\{ \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\} \\
&= \frac{2}{\sigma_e^2} \sum_{k=1}^{\varrho} \left\{ \operatorname{Re} \left[A_{diag,\omega}^*(k) R^H R A_{diag,\alpha}(k) \right] + \operatorname{Re} \left[A_{diag,\omega}^* R^H R_{\alpha} A_{diag}(k) \right] \right. \\
&\quad \left. + \operatorname{Re} \left[A_{diag}^H(k) R_{\omega}^H R A_{diag,\alpha}(k) \right] + \operatorname{Re} \left[A_{diag}^H(k) R_{\omega}^H R_{\alpha} A_{diag}(k) \right] \right\} \quad (3-36)
\end{aligned}$$

$$\begin{aligned}
\Delta &\stackrel{\Delta}{=} E \left\{ \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\} \\
&= \frac{2}{\sigma_e^2} \sum_{k=1}^{\varrho} \left\{ \operatorname{Re} \left[A_{diag,\alpha}^*(k) R^H R A_{diag,\alpha}(k) \right] + \operatorname{Re} \left[A_{diag,\alpha}^* R^H R_{\alpha} A_{diag}(k) \right] \right. \\
&\quad \left. + \operatorname{Re} \left[A_{diag}^H(k) R_{\alpha}^H R A_{diag,\alpha}(k) \right] + \operatorname{Re} \left[A_{diag}^H(k) R_{\alpha}^H R_{\alpha} A_{diag}(k) \right] \right\} \quad (3-37)
\end{aligned}$$

$$\begin{aligned}
\Omega &\stackrel{\Delta}{=} E \left\{ \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right]^T \right\} \\
&= \frac{2}{\sigma_e^2} \sum_{k=1}^{\varrho} \left\{ \operatorname{Re} \left[A_{diag,\omega}^*(k) R^H R A_{diag,\omega}(k) \right] + \operatorname{Re} \left[A_{diag,\omega}^* R^H R_{\omega} A_{diag}(k) \right] \right.
\end{aligned}$$

$$\theta = [\tilde{\alpha}_R^T(1) \quad \tilde{\alpha}_I^T(1) \quad \dots \quad \tilde{\alpha}_R^T(Q) \quad \tilde{\alpha}_I^T(Q) \quad \mathbf{d}_R^T \quad \mathbf{d}_I^T \quad \alpha^T \quad \omega^T]^T \quad (3-40)$$

Then let

$$J_{11} = \begin{bmatrix} \overline{\overline{C}} & -\overline{\overline{C}} & & & \overline{B(1)} & -\overline{B(1)} \\ \overline{C} & \overline{C} & & & \overline{B(1)} & \overline{B(1)} \\ & & \ddots & & \vdots & \vdots \\ & & & \overline{\overline{C}} & -\overline{\overline{C}} & \overline{B(Q)} & -\overline{B(Q)} \\ & & & \overline{C} & \overline{C} & \overline{B(Q)} & \overline{B(Q)} \\ \overline{B^H(1)} & -\overline{B^H(1)} & \dots & \overline{B^H(Q)} & -\overline{B^H(Q)} & \overline{F} & -\overline{F} \\ \overline{B^H(1)} & \overline{B^H(1)} & \dots & \overline{B^H(Q)} & \overline{B^H(Q)} & \overline{F} & \overline{F} \end{bmatrix} \quad (3-41)$$

$$J_{12} = \begin{bmatrix} \overline{S(1)} & \overline{W(1)} \\ \overline{S(1)} & \overline{W(1)} \\ \vdots & \vdots \\ \overline{S(Q)} & \overline{W(Q)} \\ \overline{S(Q)} & \overline{W(Q)} \\ \overline{\Gamma} & \overline{V} \\ \overline{\Gamma} & \overline{V} \end{bmatrix} \quad (3-42)$$

$$J_{21} = \begin{bmatrix} \overline{S^H(1)} & -\overline{S^H(1)} & \dots & \overline{S^H(Q)} & -\overline{S^H(Q)} & \overline{\Gamma^H} & -\overline{\Gamma^H} \\ \overline{W^H(1)} & -\overline{W^H(1)} & \dots & \overline{W^H(Q)} & -\overline{W^H(Q)} & \overline{V^H} & -\overline{V^H} \end{bmatrix} \quad (3-43)$$

$$J_{22} = \begin{bmatrix} \Delta & X \\ X^H & \Omega \end{bmatrix} \quad (3-44)$$

Then the Fisher information matrix $I_f(\theta)$ can be written as

$$I_f(\theta) = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \quad (3-45)$$

By taking the inverse of $I_f(\theta)$, we get the Cramer-Rao Lower Bound for θ

$$CRLB(\theta) = \begin{bmatrix} (J_{11} - J_{12}J_{22}^{-1}J_{21})^{-1} & -(J_{11} - J_{12}J_{22}^{-1}J_{21})^{-1}J_{12}J_{22}^{-1} \\ -(J_{22} - J_{21}J_{11}^{-1}J_{12})^{-1}J_{21}J_{11}^{-1} & (J_{22} - J_{21}J_{11}^{-1}J_{12})^{-1} \end{bmatrix} \quad (3-46)$$

For the parameters α and ω we thus have

$$CRLB^{-1}\left(\begin{bmatrix} \alpha^T & \omega^T \end{bmatrix}^T\right) = (J_{22} - J_{21}J_{11}^{-1}J_{12}) \quad (3-47a)$$

$$= \begin{bmatrix} \Delta & X \\ X^H & \Omega \end{bmatrix} - \begin{bmatrix} \Lambda & \Pi \\ \Xi & \Theta \end{bmatrix} \quad (3-47b)$$

where

$$\Lambda = \text{Re} \sum_{k=1}^{\varrho} \left\{ \Gamma^H T [\Gamma - B^H(k)C^{-1}S(k)] + \sum_{l=1}^{\varrho} S^H(l)G[S(k) - B(k)F^{-1}\Gamma] \right\} \quad (3-48)$$

$$\Xi = \text{Re} \sum_{k=1}^{\varrho} \left\{ V^H T [\Gamma - B^H(k)C^{-1}S(k)] + \sum_{l=1}^{\varrho} W^H(l)G[S(k) - B(k)F^{-1}\Gamma] \right\} \quad (3-49)$$

$$\Pi = \text{Re} \sum_{k=1}^{\varrho} \left\{ \Gamma^H T [V - B^H(k)C^{-1}W(k)] + \sum_{l=1}^{\varrho} S^H(l)G[W(k) - B(k)F^{-1}V] \right\} \quad (3-50)$$

$$\Theta = \text{Re} \sum_{k=1}^{\varrho} \left\{ V^H T [V - B^H(k)C^{-1}W(k)] + \sum_{l=1}^{\varrho} W^H(l)G[W(k) - B(k)F^{-1}V] \right\} \quad (3-51)$$

where we have defined

$$G = \left[C^{-1}\delta_{l,k} + C^{-1}B(l) \left(F - \sum_{i=1}^{\varrho} B^H(i)C^{-1}B(i) \right)^{-1} B^H(k)C^{-1} \right] \quad (3-52)$$

$$T = \left(F - \sum_{i=1}^{\varrho} B^H(i)C^{-1}B(i) \right)^{-1} \quad (3-53)$$

The proof of (3-47) through (3-51) can be found in Appendix D.

In practice the frequency parameters are often of interest, sometimes without the other parameters. From (3-47) we can obtain the Cramer-Rao Lower Bound for the frequency estimates.

$$CRLB^{-1}(\omega) = (\Omega - \Theta) - (X^H - \Xi)(\Delta - \Lambda)^{-1}(X - \Pi) \quad (3-54)$$

4. Special Case CRLB for the PAR Model

The results in the previous chapters are rather general and should produce the expected earlier derived results when considering the corresponding environments. The generalizations provided here are the colored measurement noise and the introduction of damping. If we consider white noise as a limiting form of colored noise, and an undamped signal as a limiting form of a damped signal, then present CRLB results should corroborate previous CRLB results.

4.1 Multiple Experiments in White Noise

If we consider undamped signals, i.e. $\alpha = 0$, and white measurement noise, then the partition corresponding to $\begin{bmatrix} d_R^T & d_I^T & \alpha^T \end{bmatrix}$ is removed from the parameter vector used in (3-40). Consequently the terms $B, F, S, \Gamma, V, \Delta, X$ in (3-39) do not exist, and J_{22} in (3-44) becomes equal to Ω . Equation (3-54) now reduces to

$$CRLB^{-1}(\omega) = (\Omega - \Theta) \quad (4-1)$$

Noting that $C = \bar{C} + j\bar{C}$ and $W = \bar{W} + j\bar{W}$ and using (3-26) and (3-32), equation (3-51) reduces to

$$\begin{aligned} \Theta &= \sum_{k=1}^{\varrho} \text{Re}[W^H(k)C^{-1}W(k)] \\ &= \frac{2}{\sigma_e^2} \sum_{k=1}^{\varrho} \text{Re} \left[\left(R^H R A_{diag,\omega}(k) + R^H R_\omega A_{diag}(k) \right)^H (R^H R)^{-1} \left(R^H R A_{diag,\omega}(k) + R^H R_\omega A_{diag}(k) \right) \right] \\ &= \frac{2}{\sigma_e^2} \sum_{k=1}^{\varrho} \left\{ \text{Re} \left[A_{diag,\omega}^H(k) R^H R A_{diag,\omega}(k) \right] + \text{Re} \left[A_{diag,\omega}^H(k) R^H R_\omega A_{diag}(k) \right] \right. \\ &\quad \left. + \text{Re} \left[A_{diag,\omega}^H(k) R_\omega^H R A_{diag,\omega}(k) \right] + \text{Re} \left[A_{diag}^H(k) R_\omega^H R (R^H R)^{-1} R^H R_\omega A_{diag}(k) \right] \right\} \quad (4-2) \end{aligned}$$

Substitution of (4-2) and (3-38) in (4-1) yields

$$CRLB^{-1}(\omega) = \frac{2}{\sigma_e^2} \sum_{k=1}^{\varrho} \text{Re} \left[A_{diag}^H(k) R_\omega^H R_\omega A_{diag}(k) - A_{diag}^H(k) R_\omega^H R (R^H R)^{-1} R^H R_\omega A_{diag}(k) \right]$$

$$= \frac{2}{\sigma_e^2} \sum_{k=1}^Q \text{Re} \left\{ A_{diag}^H(k) R_\omega^H \left[I - R(R^H R)^{-1} R^H \right] R_\omega A_{diag}(k) \right\} \quad (4-3)$$

Finally we get for the Cramer-Rao Lower Bound for ω :

$$CRLB(\omega) = \frac{\sigma_e^2}{2} \left[\sum_{k=1}^Q \text{Re} \left\{ A_{diag}^H(k) R_\omega^H \left[I - R(R^H R)^{-1} R^H \right] R_\omega A_{diag}(k) \right\} \right]^{-1} \quad (4-4)$$

For this special case our result in (4-4) agrees with that previously found (Equation (4-1) in [15]). The case of an undamped deterministic amplitude exponential signal in white Gaussian noise is a limiting form of the case of damped, random amplitude exponential signals in colored noise.

4.2 Single Experiment in White Noise

For a single experiment ($Q = 1$), single undamped exponential signal ($M = 1, \alpha = 0$), in white noise ($p = 0$), from (3-14), (3-15), (3-12), and (3-19) we get

$$R = \begin{bmatrix} 1 & e^{j\omega} & \dots & e^{j(N-1)\omega} \end{bmatrix}^T$$

$$A_{diag} = \tilde{\mathbf{a}}$$

$$R^H R = N$$

$$R_\omega = \begin{bmatrix} 0 & je^{j\omega} & \dots & j(N-1)e^{j(N-1)\omega} \end{bmatrix}^T$$

$$R_\omega^H R_\omega = 1^2 + 2^2 + \dots + (N-1)^2 = \frac{N(N-1)(2N-1)}{6}$$

$$R_\omega^H R = -j[1+2+\dots+(N-1)] = -j\frac{N(N-1)}{2} \quad (4-5)$$

From (4-4) we now get

$$\begin{aligned} CRLB(\omega) &= \frac{\sigma_e^2}{2} \left[\text{Re} \left\{ \tilde{\mathbf{a}}^H R_\omega^H R_\omega \tilde{\mathbf{a}} - \tilde{\mathbf{a}}^H R_\omega^H R (R^H R)^{-1} R^H R_\omega \tilde{\mathbf{a}} \right\} \right]^{-1} \\ &= \frac{6\sigma_e^2}{|\tilde{\mathbf{a}}|^2 N(N^2-1)} \\ &= \frac{6}{N(N^2-1)SNR} \end{aligned}$$

$$\approx \frac{6}{N^3 SNR} \quad \text{for large } N \quad (4-6a)$$

where we defined

$$SNR = \frac{|\tilde{a}|^2}{\sigma_e^2} \quad (4-6b)$$

This result agrees with earlier results for the single sinusoid [18] in this special case.

4.3 Single Experiment in Colored Noise

For a single experiment ($Q=1$), and a single undamped exponential signal ($M=1, \alpha=0$) in colored noise ($p \neq 0$), the partition of the parameter vector θ in (3-40) corresponding to α does not exist. Therefore $\Delta, X, X^H, \Gamma, S$ in (3-39) do not exist, and Λ, Π, Ξ in (3-47) do not exist. We now need to evaluate Ω from (3-38) and Θ from (3-51). Therefore we need

$$R = \left[e^{jp\omega} \quad e^{j(p+1)\omega} \quad \dots \quad e^{j(N-1)\omega} \right]^T \quad (4-7a)$$

$$R_\omega = \left[jpe^{jp\omega} \quad j(p+1)e^{j(p+1)\omega} \quad \dots \quad j(N-1)e^{j(N-1)\omega} \right]^T \quad (4-7b)$$

$$\begin{aligned} R_\omega^H R_\omega &= \left[p^2 + (p+1)^2 + \dots + (N-1)^2 \right] \\ &= \frac{N(N-1)(2N-1) - p(p-1)(2p-1)}{6} \triangleq S_p \end{aligned} \quad (4-7c)$$

$$\begin{aligned} R_\omega^H R &= -j \left[p + (p+1) + \dots + (N-1) \right] \\ &= -j \frac{(N+p-1)(N-p)}{2} \end{aligned} \quad (4-7d)$$

$$R^H R = (N-p) \quad (4-7e)$$

$$A_{diag} = \tilde{a} \quad (4-7f)$$

$$A_{diag,\omega} = \left[\frac{\partial \tilde{a}}{\partial \omega} \right] = j\tilde{a} \quad (4-7g)$$

From (4-7) and (3-38) we now obtain

$$\begin{aligned}
\Omega &= \frac{2}{\sigma_e^2} \operatorname{Re} \left\{ -j\tilde{\alpha}^*(N-p)j\tilde{\alpha} - j\tilde{\alpha}^*j \frac{(N+p-1)(N-p)}{2} \tilde{\alpha} \right. \\
&\quad \left. + \tilde{\alpha}^* \left(-j \frac{(N+p-1)(N-p)}{2} \right) j\tilde{\alpha} + \tilde{\alpha}^* \frac{N(N-1)(2N-1) - p(p-1)(2p-1)}{6} \tilde{\alpha} \right\} \\
&= \frac{2|\tilde{\alpha}|^2}{\sigma_e^2} \operatorname{Re} \left\{ (N-p) + 2 \frac{(N+p-1)(N-p)}{2} + \frac{N(N-1)(2N-1) - p(p-1)(2p-1)}{6} \right\} \\
&= \frac{2|\tilde{\alpha}|^2}{\sigma_e^2} \operatorname{Re} \left[(N^2 - p^2) + \frac{N(N-1)(2N-1) - p(p-1)(2p-1)}{6} \right] \tag{4-8}
\end{aligned}$$

From (3-51) in this case we obtain Θ ; the derivation is given in Appendix E.

$$\begin{aligned}
\Theta &= \frac{2}{\sigma_e^2} \operatorname{Re} \left\{ \left(A_{diag,\omega}^H R^H R + A_{diag,\omega}^H R_\omega^H R \right) \left[(R^H R)^{-1} + (R^H R)^{-1} R^H E(Y) T E(Y^H) R (R^H R)^{-1} \right] \right. \\
&\quad \left[\left(R^H R A_{diag,\omega} + R^H R_\omega A_{diag,\omega} \right) - R^H E(Y) \left(E(Y^H Y) \right)^{-1} \left(E(Y^H) R A_{diag,\omega} + E(Y^H) R_\omega A_{diag,\omega} \right) \right] \\
&\quad \left. + \left(A_{diag,\omega}^H R^H E(Y) + A_{diag,\omega}^H R_\omega^H E(Y) \right) T \times \right.
\end{aligned}$$

$$\begin{aligned}
& \times \left[E(Y^H) R A_{diag, \omega} + E(Y^H) R_{\omega} A_{diag} - E(Y^H) R (R^H R)^{-1} (R^H R A_{diag, \omega} + R^H R_{\omega} A_{diag}) \right] \\
& = \frac{2|\hat{a}|^2}{\sigma_e^2} \operatorname{Re} \left\{ \frac{(N-p)(N+p+1)^2}{4} + \frac{(N+p+1)^2 - 2(N+p-1)}{4} R^H E(Y) T E(Y^H) R + R_{\omega}^H E(Y) T E(Y^H) R_{\omega} \right. \\
& \quad - \frac{(N+p+1)}{2} R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R - \frac{(N+p+1)}{2(N-p)} R^H E(Y) T E(Y^H) R R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R \\
& \quad + j \left[\frac{(N+p+1)}{2} R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R_{\omega} + \frac{(N+p+1)}{2(N-p)} R^H E(Y) T E(Y^H) R R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R_{\omega} \right. \\
& \quad \left. \left. - \frac{(N+p-1)}{2} R_{\omega}^H E(Y) T E(Y^H) R - R^H E(Y) T E(Y^H) R_{\omega} \right] \right\} \quad (4-9)
\end{aligned}$$

where

$$T = \left[E(Y^H Y) - E(Y^H) R (R^H R)^{-1} R^H E(Y) \right]^{-1} \quad (4-10)$$

For this special case, finally the following Cramer-Rao Lower Bound for frequency estimation is obtained

$$\begin{aligned}
CRLB(\omega) = & \frac{2\sigma_e^2 S_1}{|\tilde{\alpha}|^2} \text{Re} \left\{ 1 - S_1 \left((N+p)^2 + 3 \right) R^H E(Y) T E(Y^H) R + 4R_\omega^H E(Y) T E(Y^H) R_\omega \right. \\
& - 2(N+p+1) R^H E(Y) E((Y^H Y))^{-1} E(Y^H) R - \frac{2(N+p+1)}{(N-p)} R^H E(Y) T E(Y^H) R R^H E(Y) E((Y^H Y))^{-1} E(Y^H) R \\
& + j \left[2(N+p+1) R^H E(Y) E((Y^H Y))^{-1} E(Y^H) R_\omega + \frac{2(N+p+1)}{(N-p)} R^H E(Y) T E(Y^H) R R^H E(Y) E((Y^H Y))^{-1} E(Y^H) R_\omega \right. \\
& \left. \left. - 2(N+p-1) R_\omega^H E(Y) T E(Y^H) R - 4R^H E(Y) T E(Y^H) R_\omega \right] \right\}^{-1} \quad (4-11)
\end{aligned}$$

where

$$S_1 = \frac{1}{4S_p - (N-p)(N+p-1)^2} \quad (4-12)$$

4.3.1 Colored Noise of Order 0 (White Noise)

In this case, the partition of the parameter vector θ in (3-40) corresponding to d_R, d_I, α does not exist, so that from (3-39) we note that the terms $\overline{S(\cdot)}$ and $\overline{\overline{S(\cdot)}}$ do not exist and because $p = 0$, from (4-7c) we find $S_p = \frac{N(N-1)(2N-1)}{6}$. We then obtain from (4-11)

$$\begin{aligned} CRLB(\omega) &= \frac{\sigma_e^2}{|\tilde{a}|^2} \frac{2}{4S_p - (N-p)(N+p-1)^2} \Big|_{p=0} \\ &= \frac{\sigma_e^2}{|\tilde{a}|^2} \frac{6}{N(N^2-1)} \end{aligned} \quad (4-13)$$

This result agrees with (4-6) and agrees with the previous result [18] for this special case.

4.3.2. The First Order Colored Noise Case

For this case we have from (2-15) and (3-19)

$$E(Y) = [E(y_1) \quad E(y_2) \quad \cdots \quad E(y_{N-1})]^T$$

$$E(Y^H Y) = E\left(\sum_{k=1}^{N-1} |y_k|^2\right)$$

$$R = \left[e^{jp\omega} \quad e^{j(p+1)\omega} \quad \dots \quad e^{j(N-1)\omega} \right]^T$$

$$R^H E(Y) = \sum_{k=1}^{N-1} E(y_k) e^{-jk\omega}$$

$$R_\omega = \left[jpe^{jp\omega} \quad j(p+1)e^{j(p+1)\omega} \quad \dots \quad j(N-1)e^{j(N-1)\omega} \right]^T$$

$$R_\omega^H E(Y) = -j \sum_{k=1}^{N-1} k E(y_k) e^{-jk\omega} \quad (4-14)$$

Substituting (4-14) into (4-11), yields the CRLB for first-order AR noise:

$$CRLB(\omega) = \frac{\sigma_e^2}{|\hat{a}|^2} \frac{6}{N(N-1)(N-2)} \left(\operatorname{Re} \left[1 - \frac{3T}{N(N-1)(N-2)} \times \left[4 \sum_{k=1}^{N-1} k E(y_k^*) e^{jk\omega} \sum_{l=1}^{N-1} (l+1) E(y_l) e^{-jl\omega} \right. \right. \right.$$

$$\left. \left. - 2N \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \sum_{l=1}^{N-1} (l+1) E(y_l) e^{-jl\omega} + (N+2)^2 \sum_{k=1}^{N-1} E(y_k) e^{-jk\omega} \sum_{l=1}^{N-1} E(y_l^*) e^{jl\omega} - \frac{2N+4}{N-1} \sum_{k=1}^{N-1} E(y_k) e^{-jk\omega} \times \right. \right.$$

$$\times \left[E \left(\sum_{i=1}^{N-1} |y_i|^2 \right) \right]^{-1} \sum_{l=1}^{N-1} (l+1) E(y_l^*) e^{jl\omega} \left((N-1)T^{-1} + \sum_{k=1}^{N-1} E(y_k) e^{-jk\omega} \sum_{l=1}^{N-1} E(y_l^*) e^{jl\omega} \right) \right]^{-1} \quad (4-15)$$

where

$$T = \left[E(Y^H Y) - E(Y^H) R (R^H R)^{-1} E(Y) \right]^{-1}$$

$$= \left[E \left(\sum_{i=1}^{N-1} |y_i|^2 \right) - \frac{1}{N-1} \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \sum_{l=1}^{N-1} E(y_l) e^{-jl\omega} \right]^{-1} \quad (4-16)$$

5. Simulated CRLB for PAR Model

In this chapter we present some computer simulation results and plausibility arguments towards confirmation of our bounds. The behaviors of the Cramer-Rao Lower Bound for the Pseudo-Autoregressive $\text{PAR}(M, p)$ model are illustrated with respect to the autoregressive order p , the number of signals M , the data length N , and the signal frequency separation Δf . The expectation operations in the expression for the bounds in the previous chapter have been replaced by averages computed from 50 independent noise realizations.

5.1 First Order Autoregressive Noise With a Single Signal

5.1.1 CRLB (\hat{f}) for AR(1) Noise for Different Data Length N

In this experiment a Pseudo-Autoregressive PAR(M, p) model with $p = 1$ and $M = 1$ was chosen. The first order autoregressive process has the parameter $d_1 = 0.95$, and its zero-lag covariance is unity. In this case the CRLB for the estimation of signal frequency was calculated from (4-15) for different numbers of data samples N ($N = 5, 10, 25$). For comparative purposes the CRLB for the estimation of signal frequency for the same power white noise case is also calculated, from (4-4). In Fig. 5.1a the solid line shows the spectrum of the first order autoregressive noise, while the dashed line represents the same power white noise spectrum. Frequency is given as a fraction of the sampling frequency. In Fig. 5.1b any solid line shows the CRLB function of (4-15) for data length $N=5$, with $E(\cdot)$ replaced by an estimate based on 50 independent noise realizations. The figure shows 50 such lines. In Fig. 5.1c the solid line shows the mean value of the CRLB function for estimation of signal frequency versus normalized signal frequency with data length $N=5$ based on the 50 graphs in Fig. 5.1b. Similarly, Fig. 5.1d and Fig. 5.1e present the cases for data length $N = 10$, and Fig. 5.1f and Fig. 5.1g for data length $N = 25$. From these figures we see that the CRLB for estimation of frequency gradually reduces as the number of data samples increases. In this autoregressive noise

environment, if we use the white noise assumption to estimate the signal frequency, the CRLB for estimation of frequency will be obtained as shown by the dotted line in Figures 5.1b through 5.1g. Comparing the dashed line with the solid line we observe that for the colored noise case the lower bound using the white noise assumption will generally show a pessimistic picture of the attainable variance relative to that possible with the assumption of a colored noise environment. Away from the AR peak one can do almost an order of magnitude better, even for relatively short data lengths.

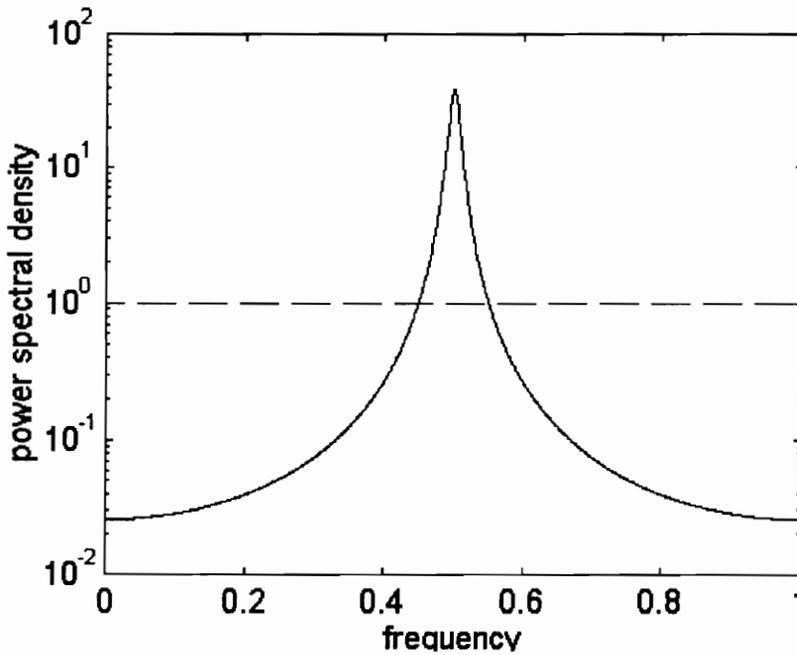


Fig. 5.1a Same power spectral densities for AR(1) noise and white noise.

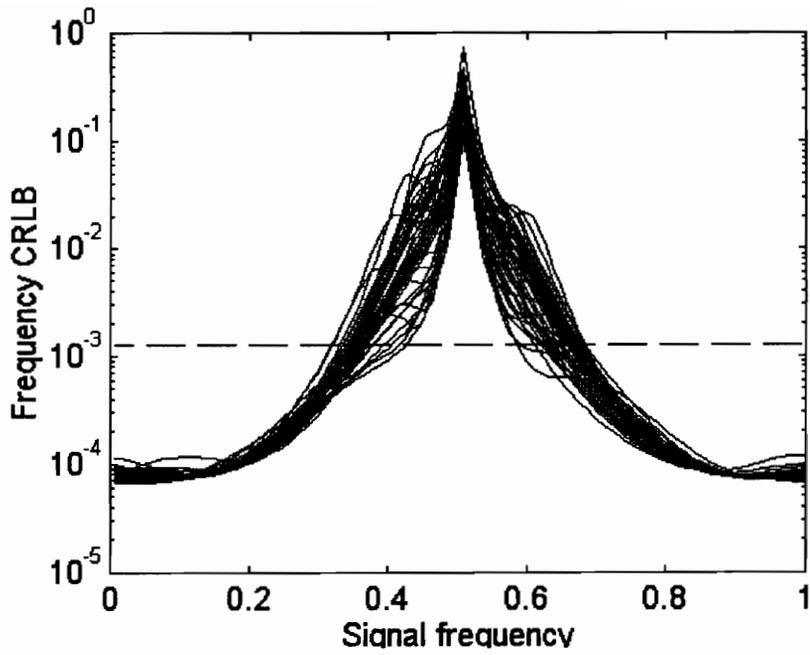


Fig. 5.1b CRLB comparison for AR(1) noise and white noise for $N=5$.

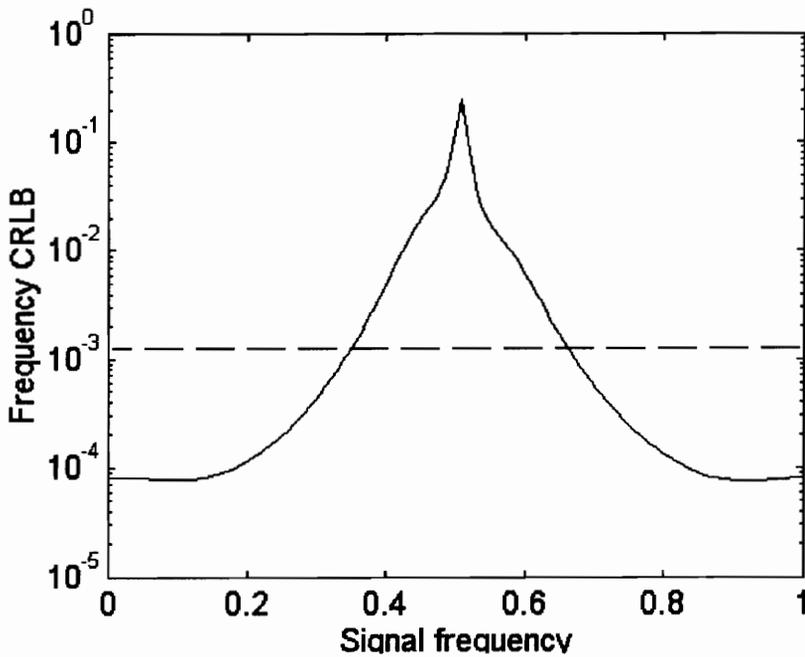


Fig. 5.1c CRLB comparison for AR(1) noise and white noise for $N=5$; mean value based on 50×50 realizations.

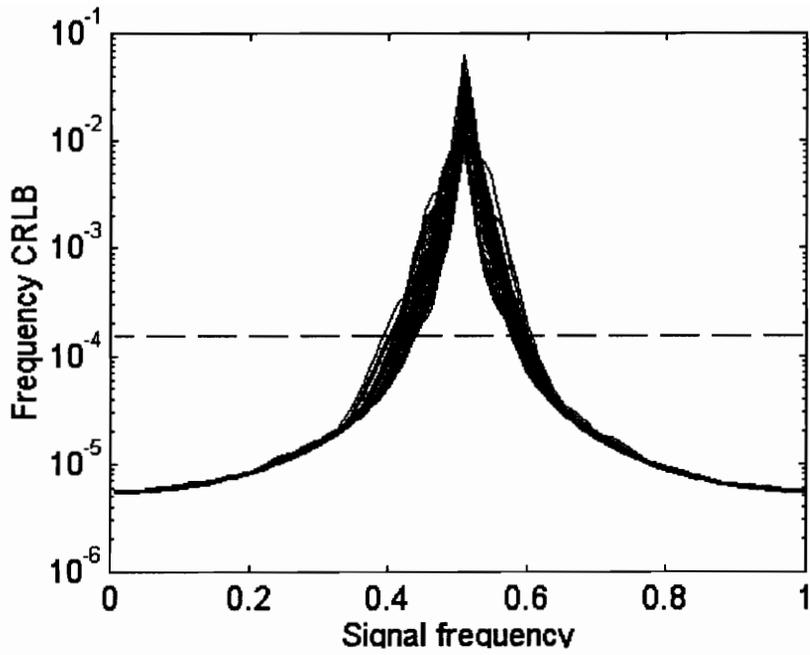


Fig. 5.1d CRLB comparison for AR(1) noise and white noise for $N=10$.

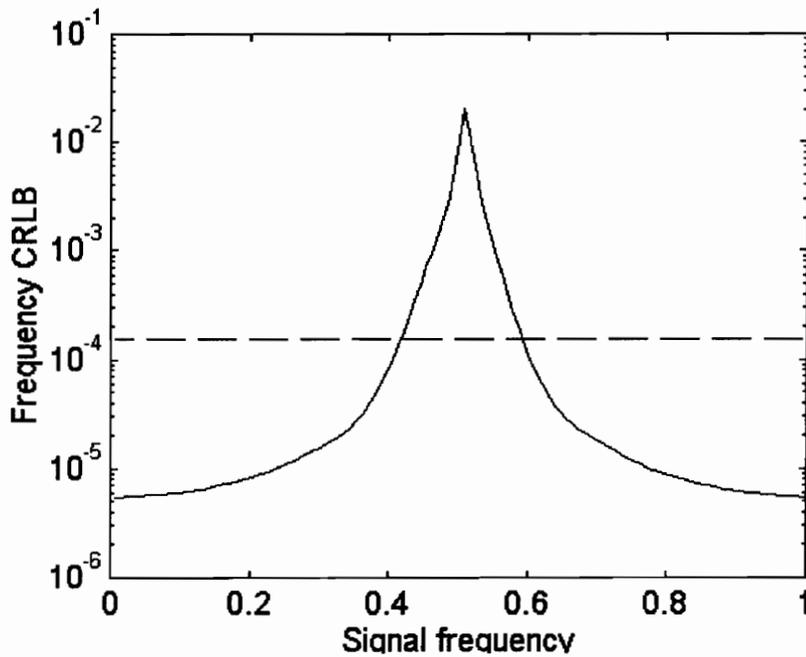


Fig. 5.1e CRLB comparison for AR(1) noise and white noise for $N=10$; mean value based on 50×50 realizations.

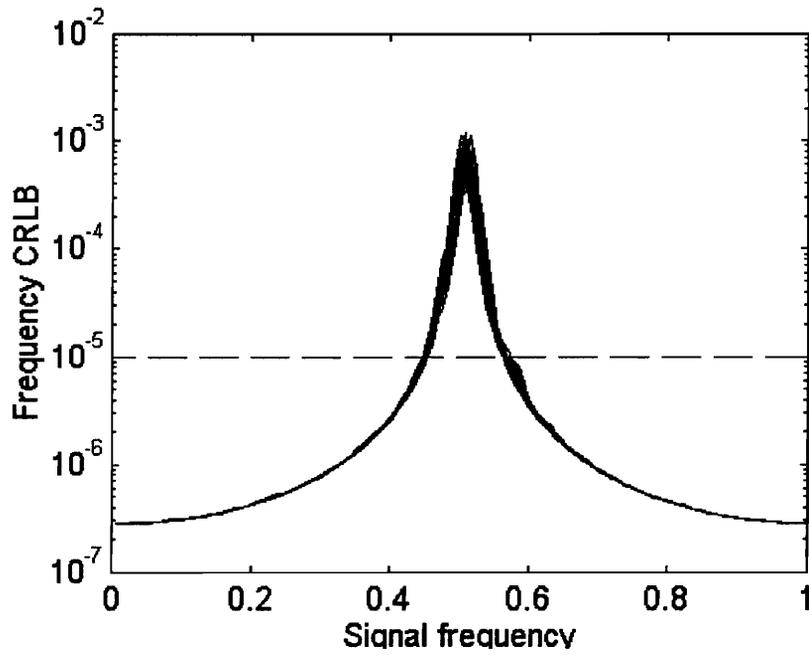


Fig. 5.1f CRLB comparison for AR(1) noise and white noise for N=25.

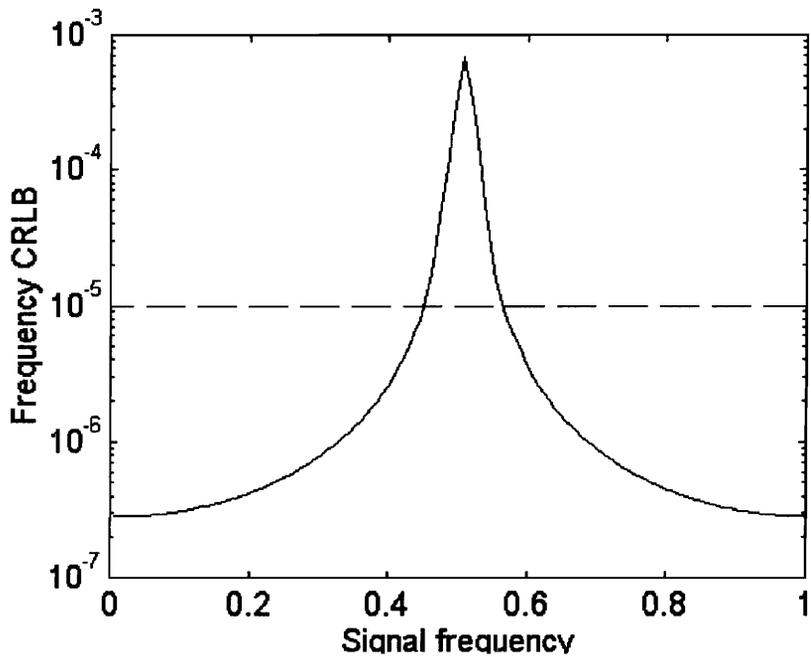


Fig. 5.1g CRLB comparison for AR(1) noise and white noise for N=25; mean value based on 50×50 realizations.

5.1.2 Different Bandwidths for Same Power AR(1) Noise

We now keep the autoregressive noise zero-lag covariance at unity for different pole radii and let the number of data samples N equal 25. In Fig. 5.2a the solid line shows the spectrum for the AR(1) process with parameter $d_1 = 0.95$, the dashed line is for $d_1 = 0.8$, the dash-dot line is for $d_1 = 0.5$, the long-dashed line is for $d_1 = 0.2$, and the short-dashed line is for the same power white noise spectrum. Fig. 5.2b shows the CRLB for estimation of frequency (simulated as in Section 5.1.1) for different d_1 ; the short-dashed line shows the CRLB for estimation of signal frequency for the same power white noise case, the dotted line shows the CRLB for estimation of signal frequency for $d_1 = 0$ case. From Fig. 5.2a, we see that when the parameter d_1 is reduced, the power spectral density of the autoregressive noise gradually closes in on the white noise case. Accordingly, Fig. 5.2b shows that when the number of data samples does not change, the CRLB for estimation of signal frequency in AR(1) noise flattens, and closes in on the CRLB for estimation of signal frequency in the white noise case as d_1 is gradually reduced.

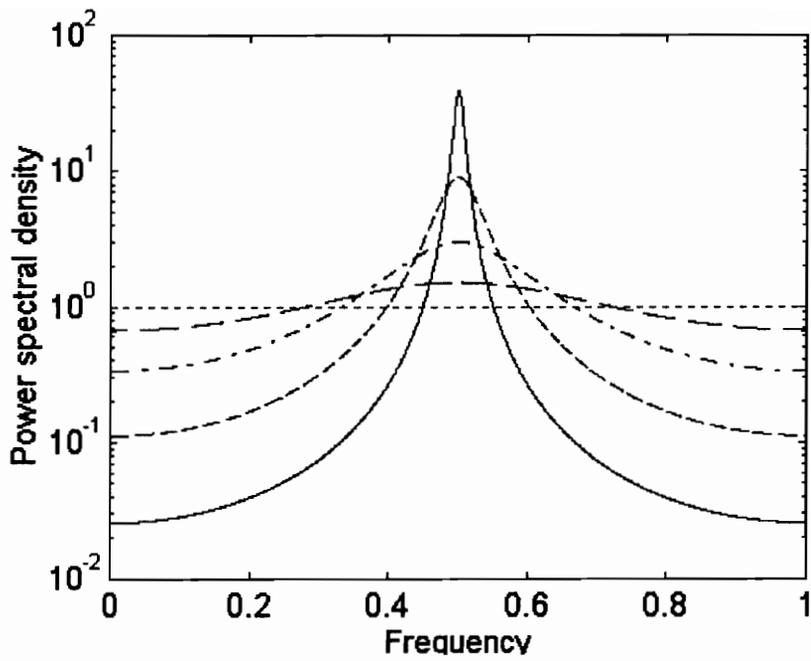


Fig. 5.2a Spectrum of AR noise with different d_1 and white noise spectrum.

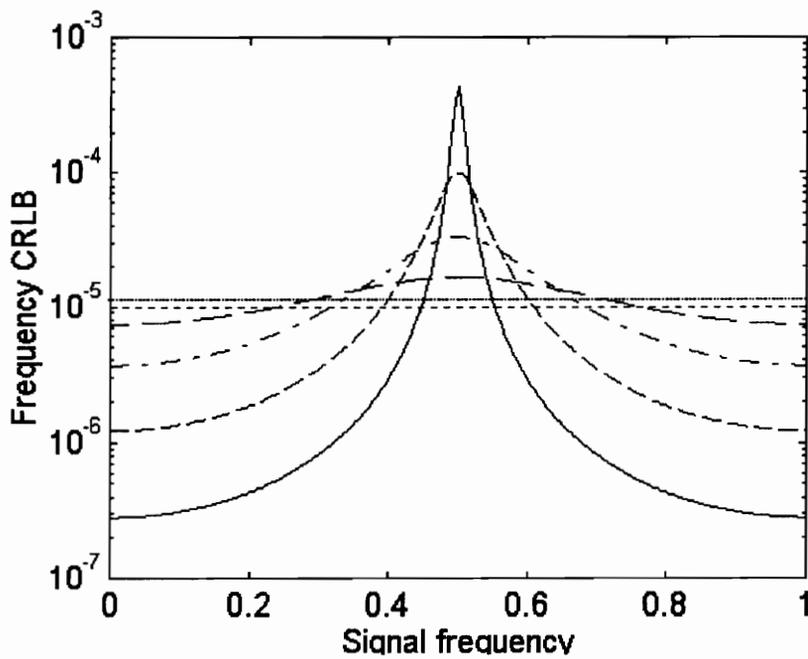


Fig. 5.2b CRLB(\hat{f}) for AR noise with different d_1 and CRLB(\hat{f}) for white noise; mean value based on 50×50 realizations.

5.2 First Order Autoregressive Noise With Two Signals

This simulation uses the $\text{PAR}(M, p)$ model with $p=1$ and $M=2$. In this experiment the first order autoregressive noise has the parameter $d_1 = 0.95$, and its zero-lag covariance equals unity. We choose $f_2 = 0.1$ as a reference frequency, and vary the frequency f_1 , where $f_1 = f_2 + \Delta f$. Fig. 5.3 illustrates the signal frequency locations with respect to the colored noise location. The CRLB for estimation of signal frequency is simulated from (3-54) for different Δf and for different numbers of data samples (simulated as in Section 5.1.1). Fig. 5.4 illustrates the CRLB for estimation of signal frequency f_1 versus the signal frequency difference Δf for various numbers of data samples N . From this figure we see that the CRLB for the estimation of signal frequency is very high for small frequency separations. This is due to the fact that signal frequency estimation becomes more difficult when the frequency difference of two signals vanishes, as is known from the white noise case [20].

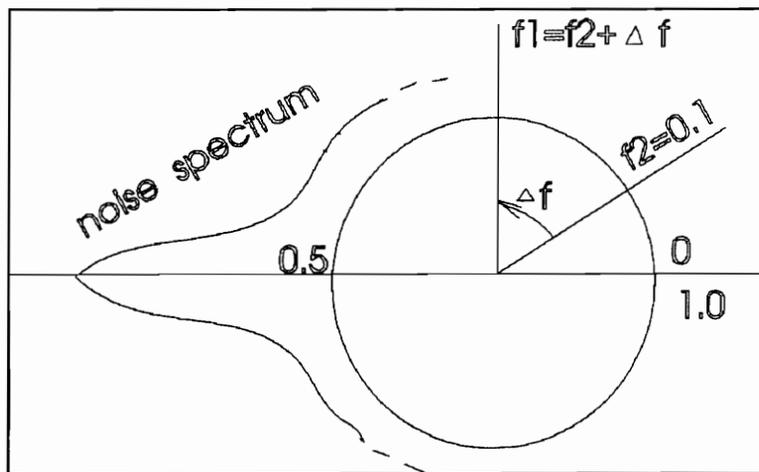


Fig. 5.3 Locations of signal frequencies and colored noise.

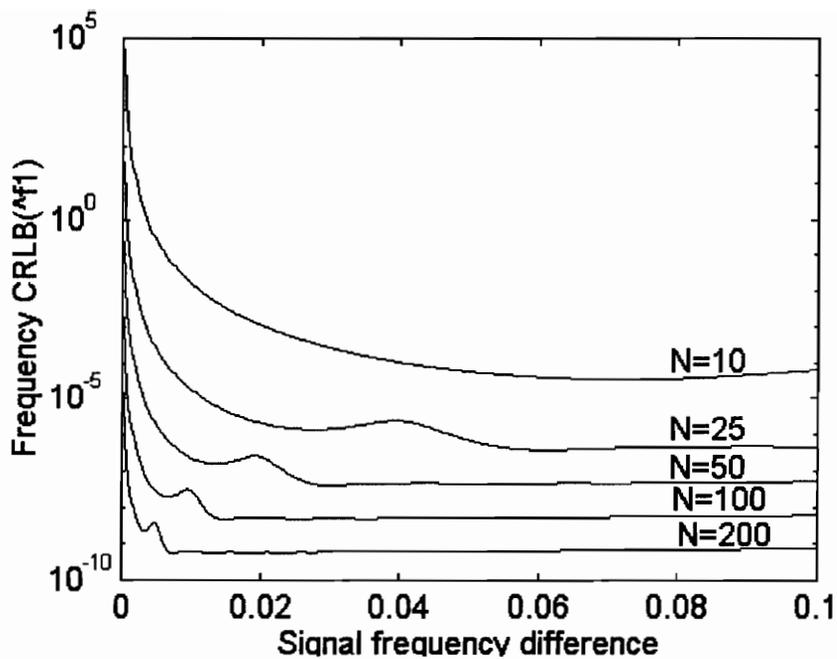


Fig. 5.4 CRLB(\hat{f}_1) versus to signal frequency difference for AR(1) noise.

Fig. 5.5 and Fig. 5.6 show the results for the entire range of frequency separation; the solid lines represent the CRLB for estimation of frequency f_1 in AR(1) noise and the dashed lines represent the CRLB for estimation of frequency f_1 in white noise. In Fig. 5.5a the solid line illustrates the CRLB for estimation of frequency f_1 in AR(1) noise with $N=10$. From Fig. 5.5b we see that when the frequency difference is small the CRLB for estimation of frequency is high, and that when the frequency difference is large enough the CRLB for estimation of signal frequency f_1 closes in on the single signal case (around $\Delta f = 0.4$ the behavior is the same as around $f = 0.5$ in Fig. 5.1e). Note that the bandwidth around $\Delta f = 0.4$ is about 0.3, which equals the bandwidth around $f = 0.5$ in Fig. 5.1e. The peak around $\Delta f = 0.4$ is just above 4×10^{-2} , whereas in Fig. 5.1e it is about 2×10^{-2} . The dashed line in Fig. 5.5b shows the CRLB for estimation of frequency f_1 in the white noise case, and we see that when the signal frequency difference is large enough, its behavior is the same as in Fig. 5.1e. Fig. 5.6 shows the same results for $N=25$. From Fig. 5.6b we see that when the frequency difference becomes large enough the CRLB for the estimation of signal frequency f_1 is relatively close to the CRLB for the single signal case shown in Fig. 5.1g (around $\Delta f = 0.4$ the behavior is the same as for $f = 0.5$ in Fig. 5.1g). Note that the bandwidth around $\Delta f = 0.4$ is about the same bandwidth around $f = 0.5$ in Fig. 5.1g. In Fig. 5.6b the peak around $\Delta f = 0.4$ is just above 8×10^{-4} , whereas in Fig. 5.1g it is less than 7×10^{-4} .

In Fig. 5.6a and 5.6b the dashed lines show the CRLB for estimation of frequency f_1 in the white noise case. From Fig. 5.6 we see that by making the colored noise assumption the bound on the variance of signal frequency estimation is about 10^{-6} , while by making the white noise assumption the bound on the variance is about 10^{-5} , an order of magnitude higher. Note that the peak value around $\Delta f = 0.4$ is closer to that in the single signal case for $N=25$ than for $N=10$. This is consistent with the CRLB approximating the single signal case CRLB when the signal frequency difference becomes large.

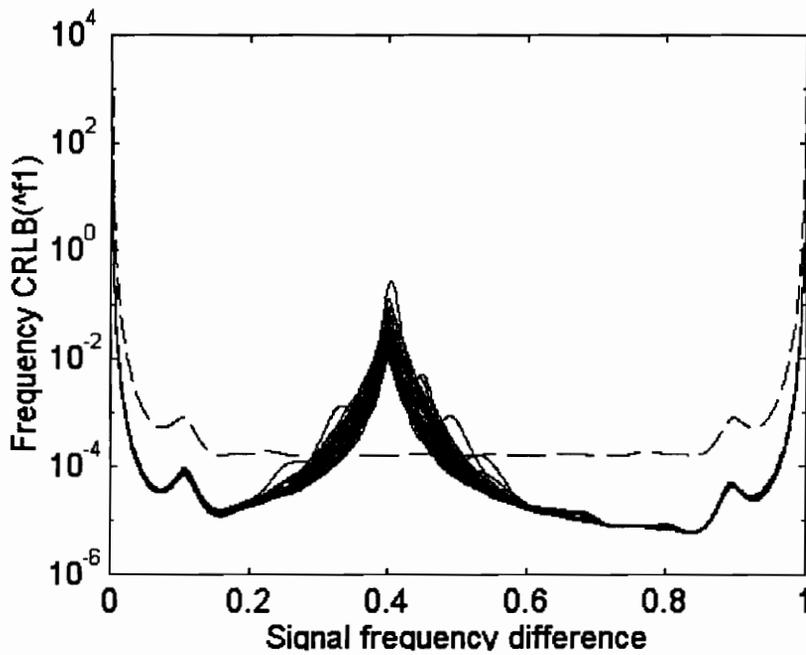


Fig. 5.5a CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for $N=10$ and 50×50 realizations.

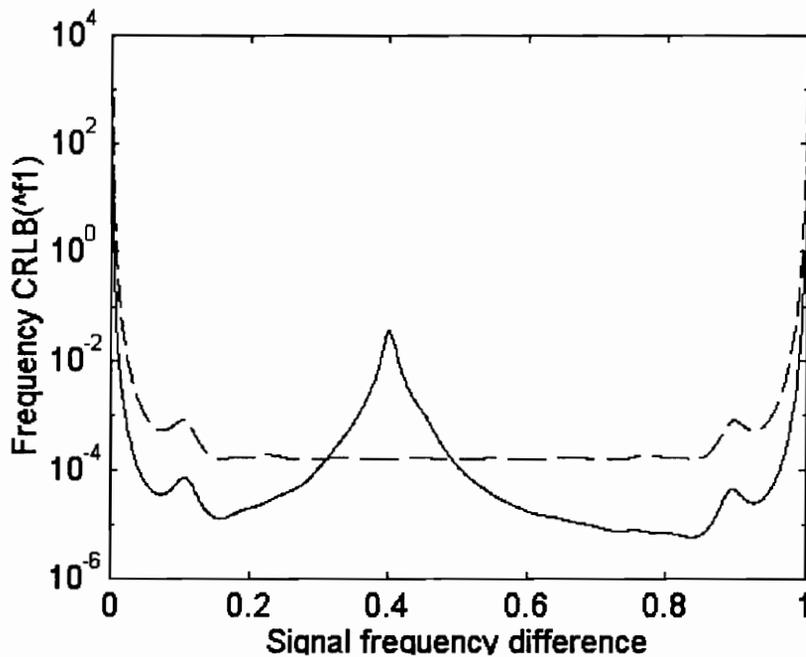


Fig. 5.5b CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for $N=10$, mean value from 50×50 realizations.

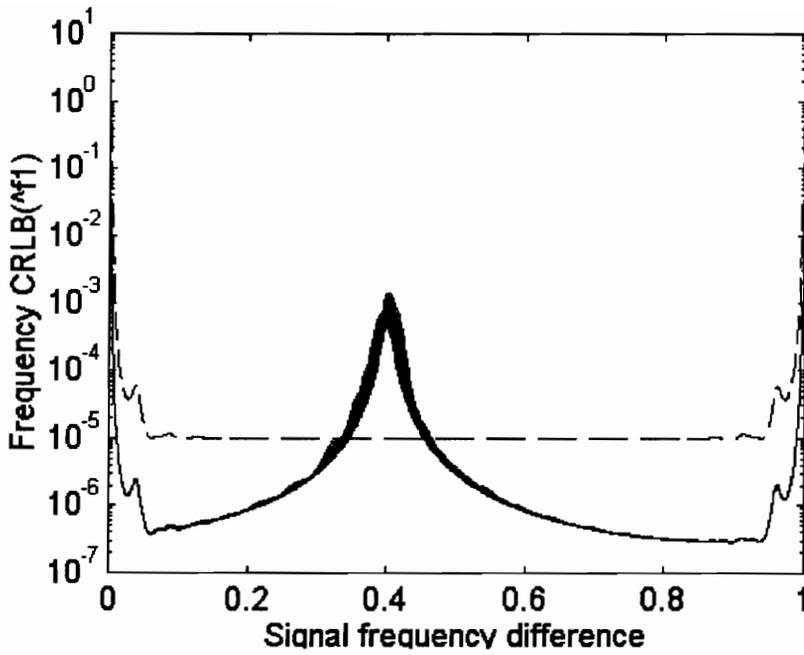


Fig. 5.6a CRLB comparison for AR(1) noise and white noise vs. signal frequency difference Δf for $N=25$ and 50×50 realizations.

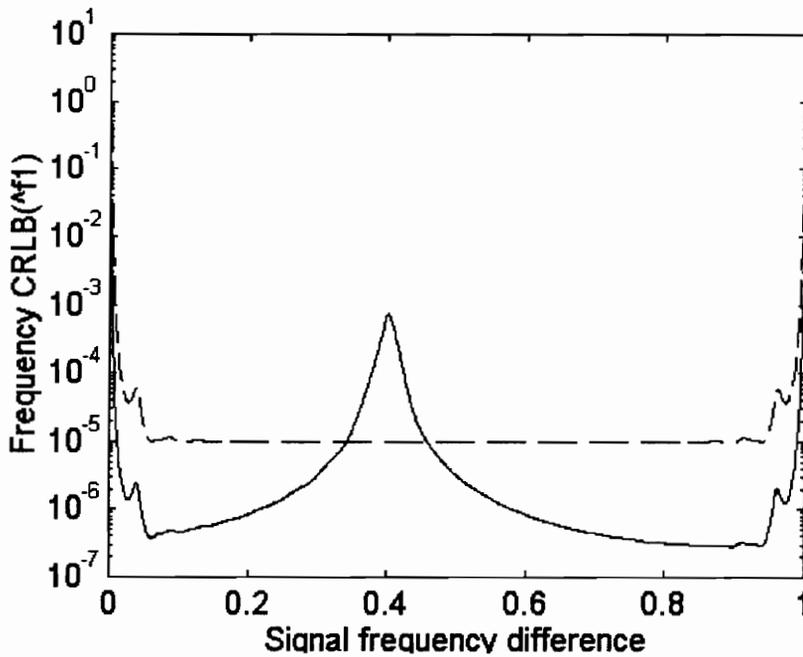


Fig. 5.6b CRLB comparison for AR(1) noise and white noise vs. Δf signal frequency difference for $N=25$; mean value from 50×50 realizations.

6. Analytical CRLB for PAR Model

6.1 PAR Model Structure Analysis

In Chapters 4 and 5 we discussed the CRLB and its simulation results for the $\text{PAR}(M, p)$ model. Now the analytical CRLB will be considered. The purpose for the analytical CRLB is that we can better understand the properties of the process at hand, if the side-effects of simulation are removed. An analytical evaluation is computationally more efficient as well.

From (2-4) we derive the Pseudo-Autoregressive (PAR) model structure, as shown in Fig. 6.1.

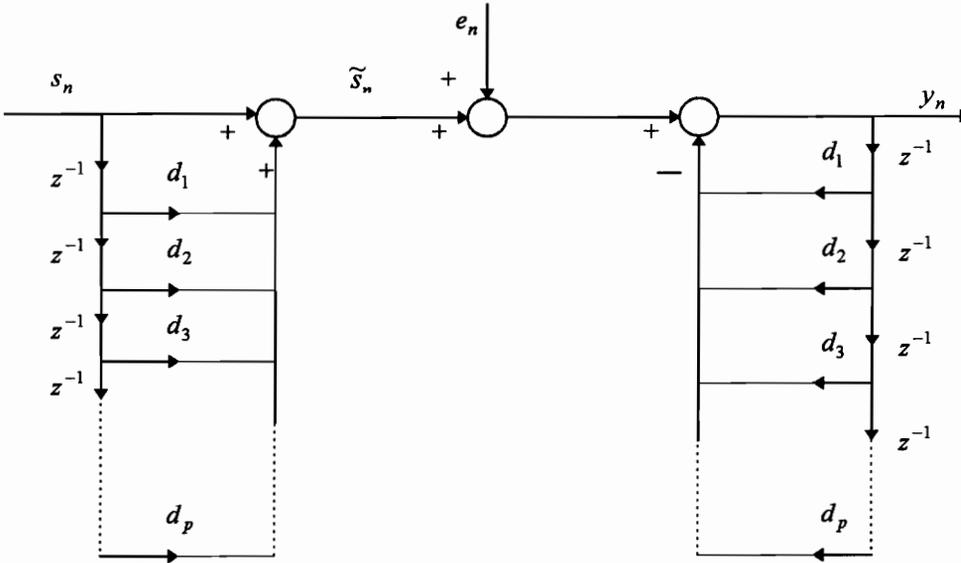


Fig. 6.1 Pseudo-Autoregressive (PAR) model structure.

From s_n to \tilde{s}_n we have the transfer function $H_1(z)$

$$\begin{aligned}
 H_1(z) &= \frac{\tilde{S}(z)}{S(z)} \\
 &= 1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_p z^{-p}
 \end{aligned} \tag{6-1}$$

From (6-1) the signal sequence \tilde{s}_n can be obtained as follows

6. Analytical CRLB for PAR Model

$$\tilde{s}_n = s_n + d_1 s_{n-1} + d_2 s_{n-2} + \dots + d_p s_{n-p} \quad (6-2)$$

Now, letting the sum of \tilde{s}_n and e_n be the input sequence, the output sequence y_n can be found through the transfer function $H_2(z)$.

$$H_2(z) = \frac{Y(z)}{\tilde{S}(z) + E(z)} \quad (6-3)$$

where $E(z)$ is the Z-transform of the input sequence e_n .

From Fig. 6.1 $H_1(z)$ and $H_2(z)$ have related structures, in particular

$$H_2(z) = \frac{1}{H_1(z)} \quad (6-4)$$

Using the transform function $H_1(z)$ and $H_2(z)$ Fig. 6.1 can be simplified as Fig. 6.2.

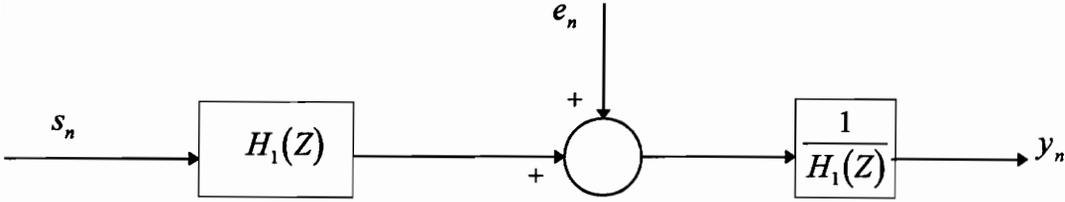


Fig. 6.2 Simplified Pseudo-Autoregressive (PAR) model structure.

By using the superposition theorem, from Fig. 6.2, we obtain the following:

(1). If there is no noise, i.e., $e_n = 0$, and there is only signal input to the system, the output is the same as the input signal. The effects of the filters $H_1(z)$ and $H_2(z) = \frac{1}{H_1(z)}$ cancel each other out. We use y_n^s to represent the output signal in this special case.

(2). If there is no signal, i.e., $s_n = 0$, and there is only the white noise sequence e_n as the input driving component, then the output is an autoregressive process. In this case the input white noise sequence is filtered by an all-pole filter. We use y_n^e to represent the output in this special case.

From the superposition theorem the system output can now be written as

$$y_n = y_n^s + y_n^e \quad (6-5)$$

In equation (4-15) we have the terms $E(Y)$ and terms $E(Y^H Y)$, where the $E(\cdot)$ represents the expectation operator and the matrix Y is defined in equation (2-15). By using (6-5) we find

$$\begin{aligned} E(Y) &= E(Y^e + Y^s) \\ &= E(Y^s) \\ &= Y^s \end{aligned} \quad (6-6a)$$

where

$$Y^s = \begin{bmatrix} y_p^s & y_{p-1}^s & \cdots & y_1^s \\ y_{p+1}^s & y_p^s & \cdots & y_2^s \\ \vdots & \vdots & \cdots & \vdots \\ y_{N-1}^s & y_{N-2}^s & \cdots & y_{N-p}^s \end{bmatrix} \quad (6-6b)$$

$$E(Y^e) = 0 \quad (6-6c)$$

Note that we use the fact that the white noise has an expected value of zero.

For the $E(Y^H Y)$ term we investigate its general (ij)-th element, i.e., the i-th row and the j-th column element of $E(Y^H Y)$.

$$\begin{aligned} E(Y^H Y)_{ij} &= E\left(\sum_{l=1}^{N-p} (y_{p+l-i}^* y_{p+l-j})\right) \\ &= E\left(\sum_{l=1}^{N-p} [(y_{p+l-i}^s + y_{p+l-i}^e)^* (y_{p+l-j}^s + y_{p+l-j}^e)]\right) \\ &= E\left(\sum_{l=1}^{N-p} [(y_{p+l-i}^s)^* y_{p+l-j}^s + (y_{p+l-i}^s)^* y_{p+l-j}^e + (y_{p+l-i}^e)^* y_{p+l-j}^s + (y_{p+l-i}^e)^* y_{p+l-j}^e]\right) \end{aligned} \quad (6-7)$$

The term y^s is the signal term, and is a deterministic term for deterministic signals.

Because y^e is an autoregressive process driven by a zero mean white noise process, its expected value equals zero,

$$E(Y^e) = 0 \quad (6-8)$$

and its autocorrelation function is

$$E\left(\left(y_{p+l-i}^e\right)^* y_{p+l-j}^e\right) = r_{y^e y^e}(i-j) \quad (6-9)$$

By using (6-8) and (6-9), equation (6-7) can be written as

$$\begin{aligned} E\left(Y^H Y\right)_{ij} &= \sum_{l=1}^{N-p} \left\{ E\left(\left(y_{p+l-i}^s\right)^* y_{p+l-j}^s\right) + E\left(\left(y_{p+l-i}^e\right)^* y_{p+l-j}^e\right) \right\} \\ &= \sum_{l=1}^{N-p} \left[\left(y_{p+l-i}^s\right)^* y_{p+l-j}^s \right] + \sum_{l=1}^{N-p} r_{y^e y^e}(i-j) \end{aligned} \quad (6-10)$$

The relationship between the parameters of an autoregressive process and its autocorrelation function follows the Yule-Walker equations [20].

$$r_{y^e y^e}(k) = \begin{cases} -\sum_{l=1}^p d(l) r_{y^e y^e}(k-l) & \text{for } k \geq 1 \\ -\sum_{l=1}^p d(l) r_{y^e y^e}(-l) + \sigma_e^2 & \text{for } k = 0 \end{cases} \quad (6-11)$$

Equation (6-11) shows the nonlinear relationship between the parameters of the AR process and the autocorrelation function. However, when given the autocorrelation function the AR parameters can be found by solving a system of linear equations from the Yule-Walker equation. When given the AR parameters we can find the autocorrelation function [58].

6.2 Analytical CRLB for Autoregressive Colored Noise Case

From (4-11), and using (6-5), (6-10), (6-11), we obtain the CRLB for the PAR model in the autoregressive colored noise case.

$$\begin{aligned}
 CRLB(\omega) = & \frac{2\sigma_e^2 S_1}{|\tilde{a}|^2} \operatorname{Re} \left\{ 1 - S_1 \left[\left((N+p)^2 + 3 \right) R^H Y^s T(Y^s)^H R + 4R_\omega^H Y^s T(Y^s)^H R_\omega \right. \right. \\
 & - 2(N+p+1)R^H Y^s \left(E(Y^H Y) \right)^{-1} (Y^s) R \\
 & \left. \left. - \frac{2(N+p+1)}{(N-p)} R^H Y^s T(Y^s)^H R R^H Y^s \left(E(Y^H Y) \right)^{-1} (Y^s) R \right. \right. \\
 & \left. \left. + j \left[2(N+p+1)R^H Y^s \left(E(Y^H Y) \right)^{-1} (Y^s) R_\omega \right. \right. \right.
 \end{aligned}$$

$$\begin{aligned}
& + \frac{2(N+p+1)}{(N-p)} R^H Y^s T(Y^s)^H R R^H Y^s (E(Y^H Y))^{-1} (Y^s) R_{\omega} \\
& - 2(N+p-1) R_{\omega}^H Y^s T(Y^s)^H R - 4 R^H Y^s T(Y^s)^H R_{\omega} \quad \left. \right\}^{-1} \quad (6-12a)
\end{aligned}$$

where

$$S_1 = \frac{1}{4S_p - (N-p)(N+p-1)^2} \quad (6-12b)$$

$$T = \left[E(Y^H Y) - (Y^s) R (R^H R)^{-1} R^H Y^s \right]^{-1} \quad (6-12c)$$

$$E(Y^H Y)_{ij} = \sum_{l=1}^{N-p} \left[(y_{p+l-i}^s)^* y_{p+l-j}^s \right] + \sum_{l=1}^{N-p} r_{y^s y^s} (i-j) \quad (6-12d)$$

Note that (6-12d) represents the ij -th element of the matrix $E(Y^H Y)$.

6.3 Analytical CRLB Evaluation Results

In this section we present some computer evaluation results and comparisons towards confirmation of our analytical bounds. The behavior of the analytical Cramer-Rao Lower Bound for the Pseudo-Autoregressive $\text{PAR}(M, p)$ model are illustrated with respect to the autoregressive order p , the number of signal M , the sample data length N , and the signal frequency separation Δf .

6.3.1 First Order Autoregressive Noise With Single Signal

In this section we evaluate the analytical CRLB for estimation of signal frequency for the first order autoregressive noise for different data length N . In this simulation a Pseudo-Autoregressive $\text{PAR}(M, p)$ model with $p = 1$ and $M = 1$ was chosen. As in Chapter 5 the first order autoregressive noise process has the parameter $d_1 = 0.95$, and its zero-lag covariance is unity. In this case the analytical CRLB for estimation of signal frequency was calculated from (6-12a) to (6-12d) with different numbers of data samples N ($N = 10$ and $N = 25$). For comparative purposes the CRLB for the estimation of signal frequency for the same power white noise case is also calculated, from (4-4). The spectrum of the first order autoregressive noise was shown in Fig. 5.1a as the solid line, while the dashed line represented the same power white noise spectrum. Frequency is

given as a fraction of sampling frequency. In Fig. 6.3 and Fig. 6.4 the solid lines show the analytical CRLB for the estimation of signal frequency in the first order autoregressive noise with a single signal, whereas the dashed lines show the CRLB for the estimation of signal frequency in the same power white noise case. In Fig. 6.3 the analytical CRLB for the estimation of signal frequency is for sample data length $N = 10$. From Fig. 6.3 we observe that the peak is about 8.5×10^{-3} and the analytical CRLB for estimation of signal frequency has a bandwidth almost the same as seen in the simulation results shown in Fig. 5.1e. In Fig. 6.4 the analytical CRLB for estimation of signal frequency is for sample data length $N = 25$, and the noise is the same as used for generating Fig. 6.3. From Fig. 6.4 we observe that the peak is about 4×10^{-4} , and the analytical CRLB for estimation of signal frequency has the same bandwidth as the statistical simulation results shown in Fig. 5.1g.

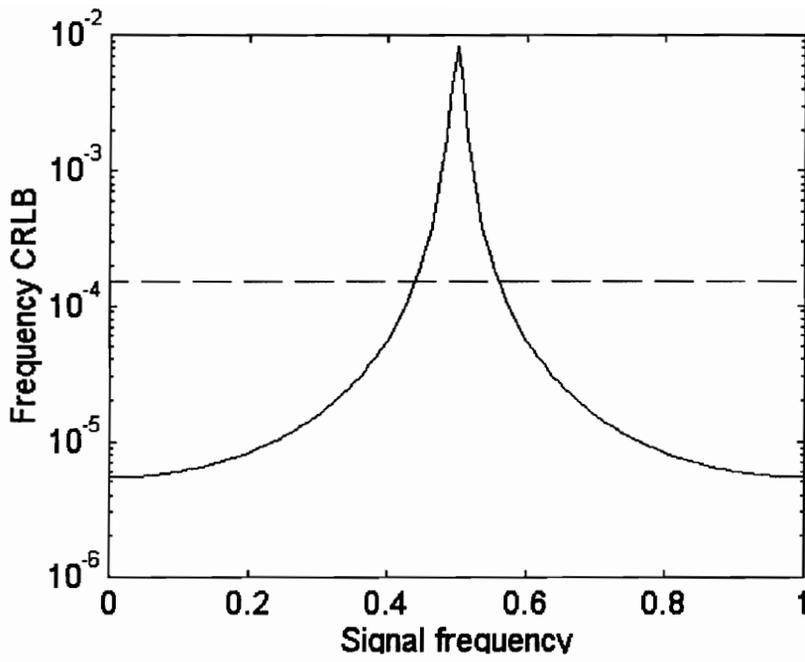


Fig. 6.3 Analytical CRLB comparison for AR(1) noise and white noise for $N = 10$.

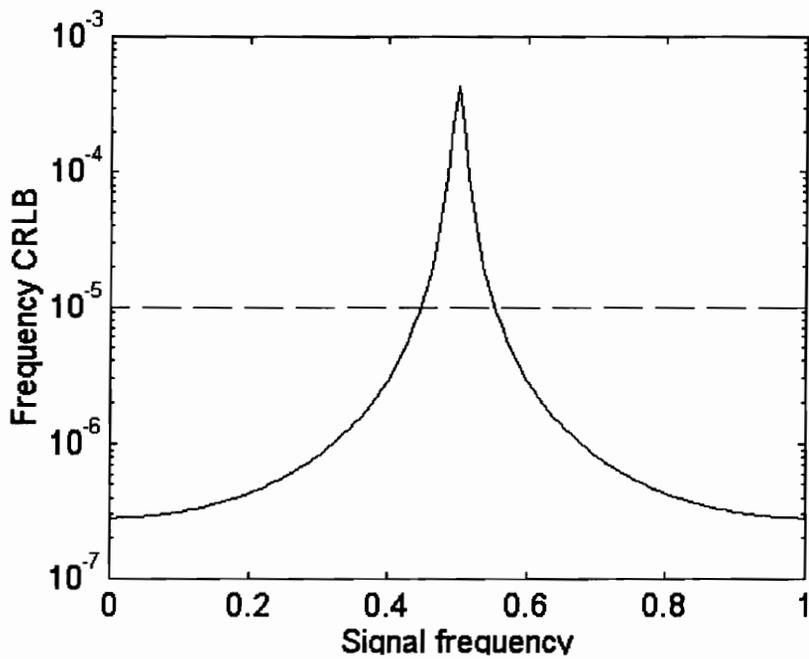


Fig. 6.4 Analytical CRLB comparison for AR(1) noise and white noise for $N = 25$.

6.3.2 First Order Autoregressive Noise With Two Signals

This evaluation uses the $\text{PAR}(M, p)$ model with $p = 1$ and $M = 2$. In this experiment the first order autoregressive noise has the parameter $d_1 = 0.95$, and its zero-lag covariance equals unity. As in the simulations, we choose the signal frequency $f_2 = 0.1$ as a reference frequency, and vary the frequency f_1 , where $f_1 = f_2 + \Delta f$. Fig. 5.3 illustrated the signal frequency locations with respect to the colored noise location. The analytical CRLB for estimation of frequency is calculated from (6-12) for different Δf and for different numbers of data samples (by simulation as in Section 5.2). Fig. 6.5 and Fig. 6.6 show the results for the entire range of frequency separation; the solid lines represent the analytical CRLB for estimation of frequency f_1 in AR(1) noise and the dashed lines represent the CRLB for estimation of frequency f_1 in white noise. In Fig. 6.5 the solid line illustrates the analytical CRLB for estimation of frequency f_1 in AR(1) noise with $N = 10$. From Fig. 6.5 we see that when the frequency difference is small the CRLB for estimation of frequency is high, and that when the frequency separation is large enough the CRLB for estimation of signal frequency f_1 closes in on the single signal case (around $\Delta f = 0.4$ the behavior is the same as around $f = 0.5$ in Fig. 6.3. Note that the bandwidth around $\Delta f = 0.4$ is approximately equal to the bandwidth around $f = 0.5$ in Fig. 6.3. The peak in Fig. 6.5 around $\Delta f = 0.4$ is 1×10^{-2} , whereas in Fig. 6.3 it is about

9×10^{-3} . The dashed line in Fig. 6.5 shows the CRLB for estimation frequency f_1 in white noise case, and we see that when the signal frequency difference is large enough, its behavior is the same as in Fig. 6.3. Note that the analytical CRLB results for estimation of frequency in Fig. 6.5 are almost the same as the simulation results shown in Fig. 5.5a, Fig. 5.5b. Fig. 6.6 shows similar results for $N = 25$. From Fig. 6.6 we see that when the frequency difference becomes large enough the CRLB for estimation of frequency f_1 is relatively close to the CRLB for the single signal case shown in Fig. 6.4 (around $\Delta f = 0.4$ the behavior is the same as for $f = 0.5$ in Fig. 6.4). Note that the bandwidth around $\Delta f = 0.4$ in Fig. 6.6 almost equals the bandwidth around $f = 0.5$ in Fig. 6.4. In Fig. 6.6 the peak around $\Delta f = 0.4$ is 5×10^{-4} , whereas in Fig. 6.4 it is 4.5×10^{-4} . In Fig. 6.6 the dashed line shows the CRLB for estimation of frequency f_1 in the white noise case. From Fig. 6.5 and Fig. 6.6 we see that the two signals CRLB approximates the single signal CRLB when the signal frequency difference becomes large enough, for large enough N .

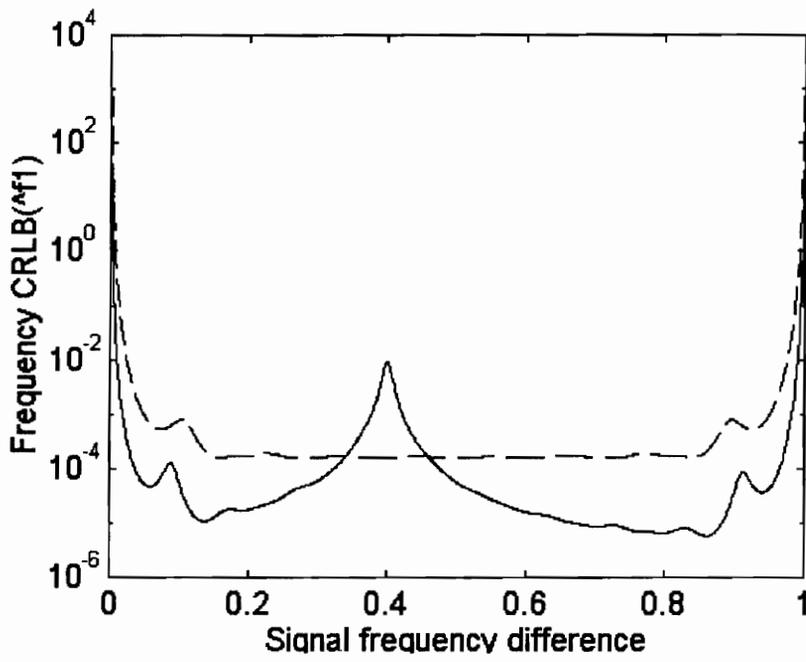


Fig. 6.5 Analytical CRLB comparison for AR(1) noise and white noise vs. Δf signal frequency difference for $N = 10$.

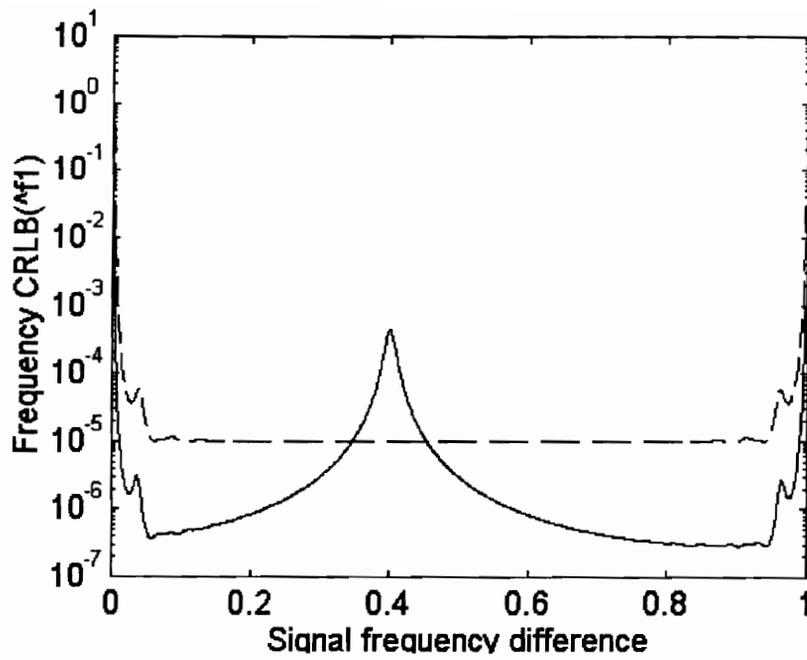


Fig. 6.6 Analytical CRLB comparison for AR(1) noise and white noise vs. Δf signal frequency difference for $N = 25$.

7. Parameter Estimation of Exponential Signals in Colored Noise

An algorithmic procedure is presented for the estimation of the parameters of exponential signals, measured in colored noise. Previous published papers on identifying sinusoids in noise have concentrated mainly on white noise disturbances. In a practical environment however, the disturbance is usually colored; sea-clutter, in a radar context, is lowpass type noise for example.

When least squares type estimates are used in the colored noise environment, this usually leads to an unacceptable bias in the estimates. In this chapter we propose an identification method named Singular Value Decomposition Bias Elimination (SVDBE). In this approach we assume that the noise can be represented well by an AR process. The parameters of this noise model are then iteratively estimated along with the exponential signal parameters, via Singular Value Decomposition (SVD) based least squares. The iteration process starts with the white noise assumption, and improves on that by allowing

the parameters in the noise model to vary away from the white noise case. A high order modal decomposition is found, and the best subset of the identified modes is selected. Simulations and comparisons with the CRLB assess the merits of the proposed algorithm.

7.1 Introduction of the Estimation Method

The traditional least squares identification approach has immediate appeal because of theoretical simplicity and ease of implementation. In colored noise environments, however, they lead to estimates with an unacceptable bias for most applications. In this chapter, which is a more complete extension of previous work [19], we propose an identification method which uses Singular Value Decomposition (SVD) to estimate and eliminate the bias in the estimated parameters. It is referred to as Singular Value Decomposition Bias Elimination (SVDBE).

The key idea in our SVDBE method is to reduce the residual, leading to an unbiased estimate. Using the SVDBE method for identifying parameters in colored noise, we can analyze the singular values and readily reduce the noise disturbing the signal, which is similar to increasing the signal-to-noise ratio (SNR). Another advantage of using the SVDBE method is its computational simplicity, which goes a long way towards

avoiding numerical instability. No matter how ill-conditioned (singular even) the system data matrix is, SVDDBE can efficiently give satisfactory results.

For clarity of presentation, we restrict attention to undamped exponential signal forms. The values of the parameters of a linear combination of exponential functions will be determined for the case of uniformly spaced samples. A short record of a data sequence y_n , $n = 1, 2, \dots, N$ is assumed to be composed of uniformly spaced samples of a sum of undamped exponential signals, x_n , and colored noise v_n .

$$y_n = x_n + v_n \quad n = 1, 2, \dots, N \quad (7-1a)$$

$$x_n = \sum_{k=1}^M a_k (c_k)^n \quad (7-1b)$$

$$N > 2M \quad (7-1c)$$

$$c_k = \exp(s_k) \quad (7-1d)$$

The parameters a_k and s_k for $k = 1, 2, \dots, M$, are unknown, complex numbers. In general the value of M is also unknown, but we will initially assume it to be known. The exponential sum signal is then the solution of a constant coefficient, linear, homogeneous

difference equation with appropriate initial conditions [51, 52], for some particular set $\{b_k\}_0^M$, where $b_0 \equiv 1$.

$$\sum_{k=0}^M b_k x_{n-k} = 0 \quad M < n \leq N \quad (7-2)$$

Alternatively we can write

$$B(q^{-1})x_n = 0 \quad M < n \leq N \quad (7-3)$$

$$B(q^{-1}) = \sum_{m=0}^M b_m q^{-m}; \quad b_0 \equiv 1 \quad (7-4)$$

where q^{-1} denotes the unit sample time delay operator.

The white noise situation has been investigated extensively [1, 7, 51, 52]. In this chapter we continue and extend the investigation of the general case in which v_n is colored noise [51].

In Section 7.2 more notation is established, and we review the fact that the Least Squares parameter estimate is a biased estimate in the colored noise situation. The SVDBE algorithm is proposed in Section 7.3, and shown to be effective for an AR

colored noise environment. The choice of the best subset out of the M determined modes is discussed, followed by some comparative simulation results for AR colored noise models.

7.2 Bias Due to Colored Noise

Operating on (7-1a) with $B(q^{-1})$ yields

$$B(q^{-1})y_n = B(q^{-1})x_n + B(q^{-1})v_n \quad (7-5)$$

Now define the noise term β_n as follows

$$\beta_n = B(q^{-1})v_n \quad (7-6)$$

Substituting (7-3), which is valid for $M < n \leq N$, in (7-5) yields

$$B(q^{-1})y_n = \beta_n \quad M < n \leq N \quad (7-7)$$

or writing out ,

$$y_n = -\sum_{k=1}^M b_k y_{n-k} + \beta_n \quad M < n \leq N \quad (7-8)$$

In matrix notation:

$$\mathbf{y} = \mathbf{Y}\mathbf{b} + \boldsymbol{\beta} \quad (7-9a)$$

$$\mathbf{Y} = \begin{bmatrix} y_M & y_{M-1} & \cdots & y_1 \\ y_{M+1} & y_M & \cdots & y_2 \\ \vdots & \vdots & \vdots & \vdots \\ y_{N-1} & y_{N-2} & \cdots & y_{N-M} \end{bmatrix} \quad (7-9b)$$

$$\mathbf{y} = \begin{bmatrix} y_{M+1} \\ y_{M+2} \\ \vdots \\ y_N \end{bmatrix}; \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_{M+1} \\ \beta_{M+2} \\ \vdots \\ \beta_N \end{bmatrix} \quad (7-9c)$$

Suppose a LS estimate $\hat{\mathbf{b}}$ of $\mathbf{b} = [b_1 \ b_2 \ \cdots \ b_M]^T$, is found from

$$\hat{\mathbf{b}} = (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{y} \quad (7-10)$$

Then the expected value of $\hat{\mathbf{b}}$ is

$$\begin{aligned}
E(\hat{\mathbf{b}}) &= E\left[(Y^T Y)^{-1} Y^T (Y\mathbf{b} + \boldsymbol{\beta})\right] \\
&= \mathbf{b} + E\left[(Y^T Y)^{-1} Y^T \boldsymbol{\beta}\right]
\end{aligned} \tag{7-11}$$

The term $E\left[(Y^T Y)^{-1} Y^T \boldsymbol{\beta}\right]$ represents the parameter bias. When β_n is white noise, then the Least Squares estimate $\hat{\mathbf{b}}$ is unbiased. In general, for the colored noise case the Least Squares estimate is biased [53]. It is this bias term that we set out to eliminate.

7.3 The SVDBE Algorithm

Assume that the colored noise residual in (7-8) satisfies an autoregressive (AR) model, that is

$$\sum_{k=0}^p d_k \beta_{n-k} = e_n \tag{7-12}$$

where p is the order of the AR noise model, and e_n is a white noise process. This model, though not our procedure, is like that for Generalized Least Squares (GLS) identification [53, 54, 55]. Equation (7-12) can be rewritten as

$$D(q^{-1})\beta_n = e_n \quad (7-13a)$$

$$D(q^{-1}) = \sum_{n=0}^p d_n q^{-n}; \quad d_0 \equiv 1 \quad (7-13b)$$

In matrix notation (7-13a) can be expressed as

$$\beta = Gd + e \quad M < n \leq N \quad (7-14a)$$

$$G = \begin{bmatrix} \beta_M & \beta_{M-1} & \cdots & \beta_{M+1-p} \\ \beta_{M+1} & \beta_M & \cdots & \beta_{M+2-p} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{N-1} & \beta_{N-2} & \cdots & \beta_{N-p} \end{bmatrix} \quad (7-14b)$$

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix}; \quad e = \begin{bmatrix} e_{M+1} \\ e_M \\ \vdots \\ e_N \end{bmatrix} \quad (7-14c)$$

Substituting (7-14a) into (7-9a), we get

$$y = [Y \quad G] \begin{bmatrix} b \\ d \end{bmatrix} + e \quad (7-15)$$

Equation (7-15) has white residuals, so the parameters can be estimated without bias, by minimizing

$$J = e^T e \quad (7-16)$$

That is

$$\begin{bmatrix} \hat{b} \\ \hat{d} \end{bmatrix} = \begin{bmatrix} Y^T Y & Y^T G \\ G^T Y & G^T G \end{bmatrix}^{-1} \begin{bmatrix} Y^T y \\ G^T y \end{bmatrix} \quad (7-17)$$

In (7-17) the matrixes $Y^T Y$ and $G^T G$ are square, so we use the partitioned matrix inverse [56]. We define thereto

$$F = (G^T G) - (G^T Y)(Y^T Y)^{-1}(Y^T G) \quad (7-18)$$

which then yields

$$\begin{bmatrix} Y^T Y & Y^T G \\ G^T Y & G^T G \end{bmatrix}^{-1} = \begin{bmatrix} (Y^T Y)^{-1} + (Y^T Y)^{-1} (Y^T G) F^{-1} (G^T Y) (Y^T Y)^{-1} & -(Y^T Y)^{-1} (Y^T G) F^{-1} \\ -F^{-1} (G^T Y) (Y^T Y)^{-1} & F^{-1} \end{bmatrix} \quad (7-19)$$

Substituting (7-19) into (7-17) yields

$$\hat{b} = (Y^T Y)^{-1} Y^T y + (Y^T Y)^{-1} (Y^T G) F^{-1} (G^T Y) (Y^T Y)^{-1} Y^T y - (Y^T Y)^{-1} (Y^T G) F^{-1} G^T y \quad (7-20)$$

$$\hat{d} = -F^{-1} (G^T Y) (Y^T Y)^{-1} Y^T y + F^{-1} G^T y \quad (7-21)$$

Substituting (7-21) into (7-20), we obtain

$$\hat{b} = (Y^T Y)^{-1} Y^T y - (Y^T Y)^{-1} (Y^T G) \hat{d} \quad (7-22)$$

The use of (7-21) and (7-20) to find \hat{d} , and then \hat{b} , is too cumbersome, because F^{-1} is needed. Here we use the simpler procedure of replacing b in (7-9a) by \hat{b} , so that

$$\hat{\beta} = y - Y\hat{b} \quad (7-23)$$

Substituting (7-23) into (7-14a), we get

$$\hat{\beta} = \hat{G}\hat{d} + e \quad (7-24)$$

Solve (7-24) for \hat{d} , using singular value decomposition [57].

$$\hat{d} = V_G \Sigma_G^{-1} U_G^T \beta \quad (7-25)$$

Now, recognizing each of the terms in (7-22) as least squares solutions to linear systems of equations, we rewrite it as follows

$$\hat{b} = V_Y \Sigma_Y^{-1} U_Y^T y - V_Y \Sigma_Y^{-1} U_Y^T (\hat{G}\hat{d}) \quad (7-26a)$$

$$= V_Y \Sigma_Y^{-1} U_Y^T (y - \hat{G}\hat{d})$$

$$= \hat{b}_0 - V_Y \Sigma_Y^{-1} U_Y^T \hat{G}\hat{d} \quad (7-26b)$$

From (7-26) we can get a consistent estimate of \hat{b} [53]. The consistent estimate \hat{b} consists of two terms. The first term corresponds to the estimation of parameters in the white noise case, and the second term is a correction due to the non-white residual β_n .

This second term in (7-26) is recognized as the estimation bias in (7-11), estimated from (7-9a) and (7-14a). We use singular value decomposition to find numerically stable parameter estimates and iteratively eliminate the estimation bias. This procedure is referred to as the Singular Value Decomposition Bias Elimination (SVDBE) method.

A summary for the SVDBE procedure is:

- 1) The initial step corresponds to the estimator for the white noise case, and is solved from (7-26) with $\hat{\mathbf{d}} \equiv \mathbf{0}$. This results in the initial parameter vector estimate $\hat{\mathbf{b}}_k$ for $k = 0$. A one time SVD of Y is performed in this step.
- 2) Use $\hat{\mathbf{b}}_k$ to generate the residual $\hat{\boldsymbol{\beta}}_k$ via (7-23) and subsequently form the estimated matrix \hat{G}_k , using (7-14b).
- 3) Use (7-25) to generate $\hat{\mathbf{d}}_k$; this involves an SVD of \hat{G}_k .
- 4) Find the new estimate $\hat{\mathbf{b}}_{k+1}$ from (7-26) by eliminating the bias in its first term, $\hat{\mathbf{b}}_0$ (the estimate for the white noise model assumption, unchanged from step 1).
- 5) Return to step 2, and repeat the process until convergence is observed.

7.4 Mode Subset Selection

To find the signal components, from a received noisy data record, is an important problem. The identification of the number of sources has practical significance in radar and sonar applications. We will consider this problem in different noise environments. First is the white noise case and then we investigate the colored noise case. The main purpose is to estimate the exponential signal components in the colored noise case, but we first investigate the white noise case. From the white noise case we get motivation for solving the signal mode subset problem in colored noise environments.

Having found an estimate for b , the roots of the polynomial $B(q^{-1})$ provide estimates for the modes $\{c_k\}_1^M$. Suppose that we have taken the initial estimate for M too high, so that we must now select an appropriate subset of the above modes. For finding the best subset of mode estimates we use the prediction error minimization criterion.

7.4.1 Mode Subset Selection in the White Noise Case

If the noise v_n in (7-1) is white noise we can use the prediction error minimization criterion to find the best subset of the modes. Under the white noise assumption, from (7-1), we get the prediction error criterion E_w for the white noise case

$$E_w = \sum_{n=1}^N \left| y_n - \sum_{k=1}^M a_k (\hat{c}_k)^n \right|^2 \quad (7-27)$$

In the investigation of the estimation accuracy of the frequencies of exponential signals observed under noisy conditions for a short record, it was found that if the prediction order M is larger than the actual number of signal components M_x , better parameter estimates are obtained. The degrees of freedom in the extra modes are used to model the noise process. When the number of signal components M_x is not known a priori, we may search for the best subset of modes in the model of (7-1b) [7]. The best subset of M modes is the one for which we best, in the least squares sense of (7-27), approximate the observed data. The following is a brief summary of this method [7].

A simple procedure is used to find the best subset is as follows:

- 1) Find an overestimated number of modes M making up the exponential signal, so that

$$M > M_x .$$

- 2) Determine the order of importance of the modes from their contribution to the summed magnitude-squared errors. The idea is to gradually delete the modes which contribute least to the reduction of error.
- 3) Calculate the approximation error from (7-27) for increasing subset size by consecutively adding the next most important mode.

Typically a rapid reduction of this error criterion is observed until M_x , and a slow reduction after that. So if M_x is not known a priori, we first choose $M_x = 1$, then calculate the corresponding error for the best subset of one mode. Next choose $M_x = 2$ and calculate the corresponding error for the best subset of two modes. This procedure is repeated until the rate of decrease of the error has become small with a further increase in the number of subset elements. The best subset size is taken to be the size where the error decrease starts to nearly vanish.

7.4.2 Best Subset Selection in the Colored Noise Case

In the colored noise case the best subset selection is more complicated than in the white noise case. Motivated by the success of the approach in the white noise case we still want to use the prediction error minimization criterion.

We first filter the colored noise data by using the estimated noise parameters $\hat{D}(q^{-1})$ and the estimated signal polynomial parameters $\hat{B}(q^{-1})$. From (7-5) we thus have

$$\hat{B}(q^{-1})y_n = \hat{B}(q^{-1})x_n + \hat{B}(q^{-1})v_n \quad (7-28)$$

and from (7-6) we know that

$$\hat{\beta}_n = \hat{B}(q^{-1})v_n \quad (7-29)$$

From (7-13a) we obtain the estimated white residual

$$\hat{D}(q^{-1})\hat{\beta}_n = \hat{e}_n \quad (7-30)$$

Pre-multiplying (7-28) by $\hat{D}(q^{-1})$ yields

$$\hat{D}(q^{-1})\hat{B}(q^{-1})y_n = \hat{D}(q^{-1})\hat{B}(q^{-1})x_n + \hat{D}(q^{-1})\hat{B}(q^{-1})v_n \quad (7-31)$$

Substituting from (7-29) and (7-30) we obtain an expression for the estimated white residual

$$\hat{e}_n = \hat{D}(q^{-1})\hat{B}(q^{-1})y_n - \hat{D}(q^{-1})\hat{B}(q^{-1})x_n \quad (7-32)$$

Note that in (7-32) the signal x_n is not known a priori, so that the estimated value, from (7-1b) is used instead.

$$\hat{x}_n = \sum_{k=1}^M a_k (\hat{c}_k)^n \quad (7-33)$$

Finally we can then use the prediction error minimization criterion E_C in the colored noise case to find the best subset.

$$E_C = \sum_{n=1}^N \left| \hat{D}(q^{-1})\hat{B}(q^{-1})y_n - \hat{D}(q^{-1})\hat{B}(q^{-1})\sum_{k=1}^M a_k (\hat{c}_k)^n \right|^2 \quad (7-34)$$

In the next section we will present some of the simulation results.

7.5 Simulation Results

Suppose the data consists of two undamped exponential signals, observed in noise.

$$y_n = a_1 \exp[j(\omega_1 n + \theta_1)] + a_2 \exp[j(\omega_2 n + \theta_2)] + v_n \quad n = 1, 2, \dots, 25 \quad (7-35)$$

where $a_1 = a_2 = 1$, $\omega_1 = 2\pi f_1$, $\omega_2 = 2\pi f_2$, $\theta_1 = \pi / 4$, $\theta_2 = 0$, $j = \sqrt{-1}$, and v_n is noise.

We evaluate the performance of the SVDBE algorithm for signal frequency estimation and the best subset selection of the signal modes for the white noise case and the AR colored noise case.

7.5.1. Signal Frequency Estimation in the White Noise Case

For comparison with previous research results we first investigate the white noise case. The signal frequencies are $f_1 = 0.52$ and $f_2 = 0.5$. SNR is defined as $10 \log_{10}(a_1^2 / (2\sigma^2))$. The variance of the real or imaginary part of the white noise is σ^2 .

For this special white noise case we then estimate the frequencies f_1 and f_2 . In order to get accurate estimates we choose $\hat{M} = [N / 2] = 12$, where $[\cdot]$ represents taking the

integer value. The above numerical choices are made to verify our results against reported ones [7]. In Table 1 the Root Mean Square Error (RMSE) for the frequency estimate \hat{f}_1 is given for different SNR. These results are based on 500 realizations and the SNR values range from 30 dB to 7 dB. These results are very close to those previously obtained by others [7].

Table 7.1

Root Mean Square Error of Frequency f_1 estimate versus Signal-to-Noise Ratio.

SNR(dB)	RMSE(\hat{f}_1)
30	3.57×10^{-4}
20	1.13×10^{-3}
15	2.01×10^{-3}
12	3.08×10^{-3}
10	4.03×10^{-3}
7	5.86×10^{-3}

7.5.2 Best Subset Selection of Modes in the White Noise Case

For this experiment we again use (7-35) to generate a 25 point data sequence. The white noise level is chosen for SNR=10 dB and SNR=30 dB. Suppose that we take as the initial estimate for the number of signal modes $\hat{M} = [N/2] = 12$. Following the prediction error minimization criterion E_w and the procedure stated in Section 7.4.1, yields the best subset selection result as in Fig. 7.1. Fig. 7.1 gives the minimum prediction error criterion, according to (7-27), as a function of the hypothesized best size M subset of the estimated 12 modes. The smallest M , after which the error stays constant, corresponds to our estimate for the number of signal components. In Fig. 7.1 the solid line represents the result for a signal-to-noise ratio equal to 10 dB, while the dashed line is the result for a signal-to-noise ratio equal to 30 dB. The estimated best mode subset size for both cases is 2, which equals the number of signal components.

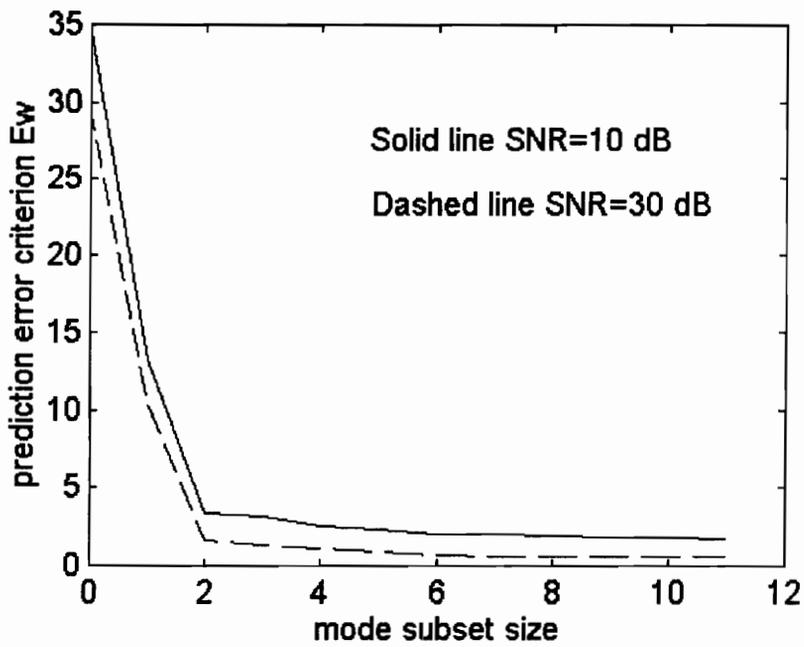


Fig. 7.1 Minimum prediction error versus subset size for white noise.

7.5.3 Signal Frequency Estimation in the AR Noise Case

In this experiment the first order autoregressive noise has the parameter $d_1 = 0.95$, and its zero-lag covariance equals unity. The power spectral density of the first order AR noise was shown as the solid line in Fig. 5.1a, where the dashed line shows the same-power white noise spectrum. We choose the signal frequencies $f_1 = 0.22$ and $f_2 = 0.2$, and the data record length $N = 25$. 100 independent realizations are used. The local signal-to-noise ratio is defined at frequency $f_0 \stackrel{\Delta}{=} (f_1 + f_2) / 2$, as

$$SNR = 10 \log_{10} \frac{|a_1|^2}{\int_{f_0-1/(2N)}^{f_0+1/(2N)} P_w(f) df} \quad \text{dB} \quad (36)$$

where the $P_w(f)$ is the power spectral density of the AR(1) noise, and a_1 is the signal amplitude. For this example the local SNR at $f_0 = 0.21$ is equal to 27.82 dB. The broadband signal-to-noise ratio is also of interest, and for this example the broadband SNR=0 dB. The broadband signal-to-noise ratio is defined as

$$SNR = 10 \log_{10} \frac{|a_1|^2}{\frac{\sigma_e^2}{1 - |d_1|^2}} \quad \text{dB} \quad (37)$$

where σ_e^2 is the driving white noise variance, and for this example $\sigma_e^2 = 0.0975$.

The SVDBE algorithm is used to estimate the signal frequencies, according to the procedure given in Section 7.3. Fig. 7.2 shows the RMSE for estimating signal frequency f_1 . The behavior for estimating signal frequency f_2 is similar. At iteration zero, the RMSE for estimating the signal frequency f_1 is equal to the RMSE under the white noise assumption. Note that an important error reduction results from using a method based on the colored noise assumption.

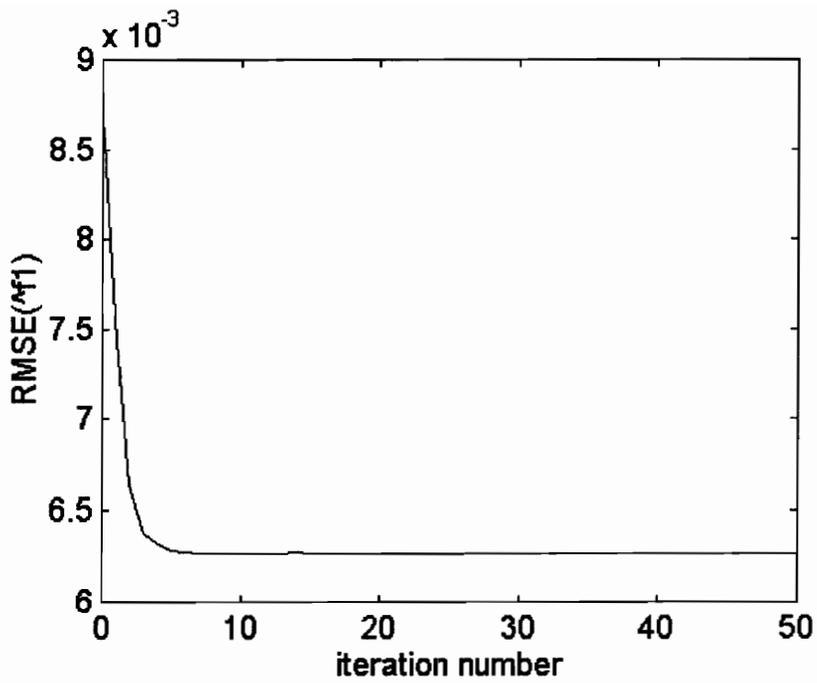


Fig. 7.2 Root mean square error for frequency estimation in AR(1) noise versus SVDBE iteration numbers.

7.5.4 Effect of Signal Frequency Difference in the AR Noise Case

Next we consider the effect of signal frequency differences for closely spaced exponential signals in autoregressive noise. Suppose the data record length is a short $N = 25$. The autoregressive noise has parameter $d_1 = 0.95$ and is the same as in Section 5.3. We chose signal frequency $f_2 = 0.1$ as the reference frequency, and varied the frequency f_1 , where $f_1 = f_2 + \Delta f$, for 100 independent realizations. At the 50-th SVDBE iteration step, the variance of \hat{f}_1 was calculated. In Fig. 7.3 the solid line shows the variance for estimation of the signal frequency f_1 versus the frequency difference. The dash-dot line shows the Cramer-Rao Lower Bound (CRLB) for exactly this example, from Chapter 6, for the estimation of signal frequency f_1 versus the frequency difference. The dashed line in Fig. 7.3 is the Cramer-Rao Lower Bound for the same-power white noise case. From Fig. 7.3 we see that at $\Delta f = 0.02$, the estimated variance for signal frequency \hat{f}_1 is about 4×10^{-5} , a value that agrees with the value 3.91×10^{-5} from Fig. 7.2 (square of 6.3×10^{-3}). We note that the SVDBE algorithm is a relatively accurate estimation algorithm which benefits from making the colored noise assumption. This holds for scenarios where the noise produces the main effect, i.e. for signal frequency differences in excess of about 0.1.

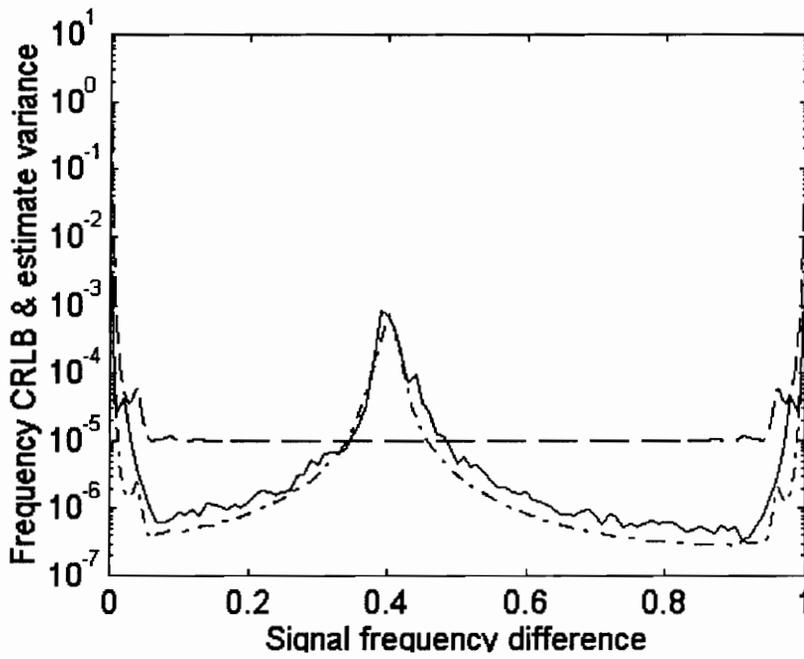


Fig. 7.3 Variance for frequency f_1 estimation and CRLB for frequency f_1 estimation for AR(1) noise and CRLB for white noise versus signal frequency difference Δf for $N = 25$.

7.5.5 Mode Subset Selection in the AR Noise Case

For the autoregressive colored noise case the best mode subset selection is obtained from the prediction error criterion E_C developed in Section 7.4.2. In this experiment we use equation (7-34) with data sequence length $N = 25$, and an autoregressive colored noise with parameter $a_1 = 0.95$. The power spectral density for this first order autoregressive colored noise and for the same power white noise is as shown in Fig. 5.1a as the solid line and the dashed line respectively. We choose $f_1 = 0.52$ and $f_2 = 0.5$. The local SNR is, at the center frequency $f_0 = 0.51$, -2.3 dB and the broadband SNR=0 dB. The initial estimate for \hat{M} equals 12; i.e. the integer part of $[N/2]$. Fig. 7.4 shows the minimum prediction error criterion E_C as a function of hypothesized subset size. Note that only until $\hat{M}_x = 2$ there is a rapid reduction of the error criterion, so that $\hat{M}_x = 2$ provides the best subset size selection. This is in agreement with the number of actual exponential signal components. For comparison we also evaluate mode subset selection according to (7-27), under the white noise hypothesis. Fig. 7.5 shows that there is no mode subset size at which the prediction error criterion reduction vanishes. Again we see that the white noise hypothesis can be detrimental to parameter estimation in the colored noise case.

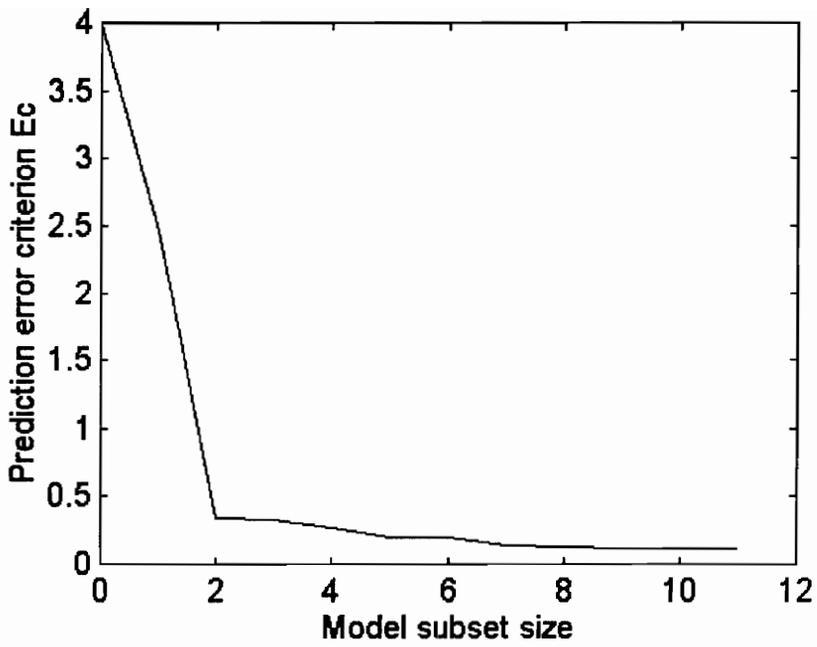


Fig. 7.4 The best subset selection for autoregressive colored noise.

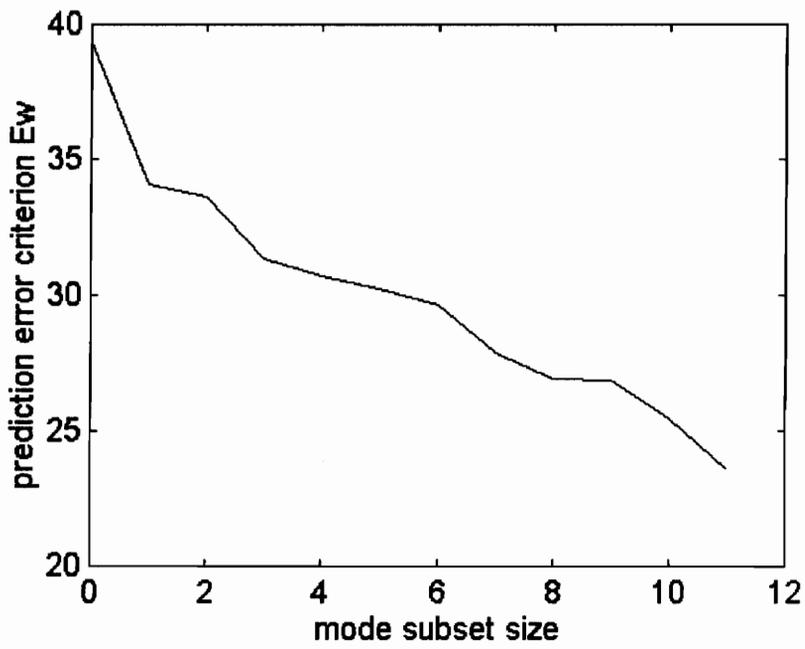


Fig. 7.5 The best subset selection for incorrect noise assumption case.

8. Conclusions

We proposed the Pseudo-Autoregressive PAR(M, p) model for the estimation of the parameters of superimposed exponential signals in colored noise, for an arbitrary number (M) of signal components and arbitrary order (p) autoregressive noise. For the various cases of exponential signals, encompassing both the time series and the uniform linear array problem in the colored noise environment, the Pseudo-Autoregressive model can be used to obtain better parameter estimates than with the white noise assumption.

The Cramer-Rao Lower Bound (CRLB) for estimating the parameters of the Pseudo-Autoregressive model is derived. This CRLB can be used to evaluate, or compare, the performance of various unbiased estimators. We have shown that our results revert to the known results for the white noise case, when viewing white noise as a limiting form of autoregressive noise. Some examples show that it is advantageous to be able to make the colored noise assumption, because - away from the AR noise peak - the

frequency estimation variance can be reduced by almost an order of magnitude for even relatively short data lengths.

We note that the estimation of parameters for exponential data observed under noisy conditions, is a difficult problem. The least squares extension of Prony's method is a simple procedure for determining the values of the parameters in a linear combination of exponential functions. Recent versions of this procedure were quite sensitive to the presence of noise, leading to poor parameter estimates under low SNR conditions. We propose a method, the SVDBE procedure, to estimate the parameters of exponential signals observed in a colored noise environment. The algorithm is investigated under the AR noise model hypothesis. Our limited simulation results are encouraging, in that they show that improvements in the estimates can result when the white noise assumption is replaced by a colored noise assumption. Another advantage of using the SVDBE algorithm is its computational simplicity and robustness towards avoiding numerical instability. No matter how ill-conditioned (singular even) the system data matrix is, the SVDBE can efficiently give satisfactory results.

As the procedure is based on estimating a relatively large number of modes, a systematic procedure for signal mode subset choice is discussed. The best mode subset selection is investigated by proposing a prediction error minimization criterion for closely spaced exponential signals in colored noise environments. Here we investigate both the white noise case and the colored noise case. For the white noise case the best subset

selection of the signal mode is the same as in previous research [7,19]. Motivated by the white noise case we proposed a prediction error minimization criterion to find the best subset of the signal modes in colored noise environments. This procedure is to use the SVDBE algorithm to find the filter parameters, filter the received data sequence and the SVDBE estimated signal components, and next use the prediction error minimization criterion to finally get the best subset selection of the signal modes. This procedure gives us a way to find the best estimates of the signal components. This criterion is simple, useful and reliable.

Appendix A

Proof of (2.11), the conditional pdf for the PAR(M,p) model.

Equation (2.4) directly yields

$$e_n = y_n + \sum_{i=1}^p d_i y_{n-i} - \sum_{k=1}^M \hat{a}_k \exp((- \alpha_k + j \omega_k) n) \quad (\text{A.1})$$

This represents a transformation from $\{y_n\}$ to $\{e_n\}$. Define

$$\mathbf{e} = \begin{bmatrix} e_{p+1} \\ e_{p+2} \\ \vdots \\ e_N \end{bmatrix}; \quad \mathbf{y} = \begin{bmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_N \end{bmatrix}; \quad \tilde{\mathbf{y}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \quad (\text{A.2})$$

We can express (A.1) in matrix notation, as

$$\mathbf{e} = \begin{bmatrix} 1 & & & & \\ d_1 & 1 & & & \mathbf{0} \\ d_2 & d_1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & d_p & \cdots & 1 \end{bmatrix} \mathbf{y} + \begin{bmatrix} d_p & d_{p-1} & d_{p-2} & \cdots & d_1 \\ & d_p & d_{p-1} & \cdots & d_2 \\ & & d_p & \cdots & d_3 \\ & & \mathbf{0} & \ddots & \vdots \\ & & & & d_p \end{bmatrix} \tilde{\mathbf{y}} - \begin{bmatrix} \sum_{k=1}^M \hat{a}_k \rho_k^{p+1} \\ \sum_{k=1}^M \hat{a}_k \rho_k^{p+2} \\ \sum_{k=1}^M \hat{a}_k \rho_k^{p+3} \\ \vdots \\ \sum_{k=1}^M \hat{a}_k \rho_k^N \end{bmatrix} \quad (\text{A.3})$$

This is

$$\mathbf{e} = \mathbf{L}\mathbf{y} + \mathbf{U}\tilde{\mathbf{y}} - \begin{bmatrix} \sum_{k=1}^M \hat{a}_k \rho_k^{p+1} \\ \sum_{k=1}^M \hat{a}_k \rho_k^{p+2} \\ \sum_{k=1}^M \hat{a}_k \rho_k^{p+3} \\ \vdots \\ \sum_{k=1}^M \hat{a}_k \rho_k^N \end{bmatrix} \quad (\text{A.4})$$

where

$$\mathbf{L} = \begin{bmatrix} 1 & & & & \\ d_1 & 1 & & & \mathbf{0} \\ d_2 & d_1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}; \quad \mathbf{U} = \begin{bmatrix} d_p & d_{p-1} & d_{p-2} & \cdots & d_1 \\ & d_p & d_{p-1} & \cdots & d_2 \\ & & d_p & \cdots & d_3 \\ & & \mathbf{0} & \ddots & \vdots \\ & & & & d_p \end{bmatrix} \quad (\text{A.5})$$

The probability density function of e , since e_n is white with variance σ_e^2 , is

$$\begin{aligned}
 p(e) &= \prod_{n=p+1}^N \frac{1}{\pi\sigma_e^2} \exp\left(-\frac{e_n^2}{\sigma_e^2}\right) \\
 &= \left(\frac{1}{\pi\sigma_e^2}\right)^{N-p} \exp\left(-\frac{1}{\sigma_e^2} \sum_{n=p+1}^N e_n^2\right)
 \end{aligned} \tag{A.6}$$

The exact maximization of (A-6) with respect to the AR parameters produces a set of highly nonlinear equations [22]. For large data records the maximization of the probability density function can be approximated by maximizing the conditional probability density function. The effect of the probability density function of the initial conditions $p(y_1, y_2, \dots, y_p; \xi)$ on the maximum likelihood estimate will be small as long as the poles are not too close to the unit circle [20, 22]. With this approximation the conditional probability density function can be maximized by using the proposed PAR(M, p) model with a suitable algorithm introduced in this dissertation.

The Jacobian of the transformation $\frac{\partial e}{\partial y}$ is just the matrix L and hence $\det(L) = 1$.

Using the formula for the probability density function of a set of transformed random variables [21], we obtain

$$\begin{aligned}
p(\mathbf{y}|\tilde{\mathbf{y}}; \xi) &= p(\mathbf{e}(\mathbf{y})) \left| \det \left(\frac{\partial \mathbf{e}}{\partial \mathbf{y}} \right) \right| = p(\mathbf{e}(\mathbf{y})) \\
&= \left(\frac{1}{\pi \sigma_e^2} \right)^{N-p} \exp \left\{ -\frac{1}{\sigma_e^2} \sum_{n=p+1}^N \left| y_n - \sum_{k=1}^M \hat{a}_k \exp(-\alpha_k n + j\omega_k n) + \sum_{i=1}^p d_i y_{n-i} \right|^2 \right\} \quad (\text{A.7})
\end{aligned}$$

Hence we proved equation (2.11).

Appendix B

1) Proof of (3-5)

From (3-2) we get

$$\begin{aligned}\frac{\partial \ln L}{\partial \mathbf{d}_R} &= -\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left\{ \frac{\partial}{\partial \mathbf{d}_R} [\mathbf{d}^H Y^H(k) \mathbf{e}(k)] + \frac{\partial}{\partial \mathbf{d}_R} [\mathbf{e}^H(k) Y(k) \mathbf{d}] \right\} \\ &= -\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left\{ Y^H(k) \mathbf{e}(k) + (\mathbf{e}^H(k) Y(k))^T \right\} \\ &= -\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left\{ Y^H(k) \mathbf{e}(k) + Y^T(k) \mathbf{e}^*(k) \right\} \\ &= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Re} [Y^H(k) \mathbf{e}(k)]\end{aligned}\tag{B.1}$$

2) Proof of (3-6)

From (3-2) we get

$$\begin{aligned}
\frac{\partial \ln L}{\partial \mathbf{d}_I} &= -\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left\{ \frac{\partial}{\partial \mathbf{d}_I} [\mathbf{d}^H Y^H(k) \mathbf{e}(k)] + \frac{\partial}{\partial \mathbf{d}_I} [\mathbf{e}^H(k) Y(k) \mathbf{d}] \right\} \\
&= -\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left\{ -jY^H(k) \mathbf{e}(k) + j(\mathbf{e}^H(k) Y(k))^T \right\} \\
&= -\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left\{ -jY^H(k) \mathbf{e}(k) + jY^T(k) \mathbf{e}^*(k) \right\} \\
&= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \text{Im}[Y^H(k) \mathbf{e}(k)] \tag{B.2}
\end{aligned}$$

3) Proof of (3-7)

From (3-2) we get

$$\begin{aligned}
\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} &= \frac{1}{\sigma_e^2} \left\{ \frac{\partial}{\partial \tilde{\mathbf{a}}_R(k)} [\tilde{\mathbf{a}}^H(k) R^H \mathbf{e}(k)] + \frac{\partial}{\partial \tilde{\mathbf{a}}_R(k)} [\mathbf{e}^H(k) R \tilde{\mathbf{a}}(k)] \right\} \\
&= \frac{1}{\sigma_e^2} \left\{ R^H \mathbf{e}(k) + (\mathbf{e}^H(k) R)^T \right\}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sigma_e^2} \{R^H \mathbf{e}(k) + R^T \mathbf{e}^*(k)\} \\
&= \frac{2}{\sigma_e^2} \operatorname{Re}[R^H \mathbf{e}(k)] \quad k = 1, 2, \dots, Q \quad (\text{B.3})
\end{aligned}$$

4) Proof of (3-8)

From (3-2) we get

$$\begin{aligned}
\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_1(k)} &= \frac{1}{\sigma_e^2} \left\{ \frac{\partial}{\partial \tilde{\mathbf{a}}_1(k)} [\tilde{\mathbf{a}}(k) R^H \mathbf{e}(k)] + \frac{\partial}{\partial \tilde{\mathbf{a}}_1(k)} [\mathbf{e}^H(k) R \tilde{\mathbf{a}}(k)] \right\} \\
&= \frac{1}{\sigma_e^2} \left\{ -j R^H \mathbf{e}(k) + j (\mathbf{e}^H(k) R)^T \right\} \\
&= \frac{1}{\sigma_e^2} \left\{ -j R^H \mathbf{e}(k) + j R^T \mathbf{e}^*(k) \right\} \\
&= \frac{2}{\sigma_e^2} \operatorname{Im}[R^H \mathbf{e}(k)] \quad k = 1, 2, \dots, Q \quad (\text{B.4})
\end{aligned}$$

5) Proof of (3-9)

From (3-2) we get

$$\begin{aligned}
\frac{\partial \ln L}{\partial \alpha_1} &= -\frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \left[\frac{\partial}{\partial \alpha_1} (y(i) - R\tilde{a}(i) + Yd)^H \right] e(i) + e^H(i) \left[\frac{\partial}{\partial \alpha_1} (y(i) - R\tilde{a}(i) + Yd) \right] \right\} \\
&= -\frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \left[-\frac{\partial \tilde{a}^H(i)}{\partial \alpha_1} R^H - \tilde{a}^H(i) \frac{\partial R^H}{\partial \alpha_1} \right] e(i) + e^H(i) \left[-\frac{\partial R}{\partial \alpha_1} \tilde{a}(i) - R \frac{\partial \tilde{a}(i)}{\partial \alpha_1} \right] \right\} \\
&= \frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \frac{\partial \tilde{a}^H(i)}{\partial \alpha_1} R^H e(i) + \tilde{a}^H(i) \frac{\partial R^H}{\partial \alpha_1} e(i) + e^H(i) \frac{\partial R}{\partial \alpha_1} \tilde{a}(i) + e^H(i) R \frac{\partial \tilde{a}(i)}{\partial \alpha_1} \right\} \\
&= \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \operatorname{Re} \left[\frac{\partial \tilde{a}^H(i)}{\partial \alpha_1} R^H e(i) + \tilde{a}^H(i) \frac{\partial R^H}{\partial \alpha_1} e(i) \right] \right\} \quad l=1, 2, \dots, M \quad (\text{B.5})
\end{aligned}$$

when $l=1$, then

$$\frac{\partial \ln L}{\partial \alpha_1} = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \operatorname{Re} \left(\left[\frac{\partial \tilde{a}_1^*(i)}{\partial \alpha_1} \quad 0 \quad \dots \quad 0 \right] \begin{bmatrix} \rho_1^p & \rho_1^{p+1} & \dots & \rho_1^{N-1} \\ \rho_2^p & \rho_2^{p+1} & \dots & \rho_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_M^p & \rho_M^{p+1} & \dots & \rho_M^{N-1} \end{bmatrix}^* e(i) \right) \right\}$$

$$\begin{aligned}
& + \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left[\tilde{a}_1^*(i) \quad \tilde{a}_2^*(i) \quad \dots \quad \tilde{a}_M^*(i) \right] \begin{bmatrix} \frac{\partial \rho_1^p}{\partial \alpha_1} & \frac{\partial \rho_1^{p+1}}{\partial \alpha_1} & \dots & \frac{\partial \rho_1^{N-1}}{\partial \alpha_1} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}^* \mathbf{e}(i) \right\} \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left[\frac{\partial \tilde{a}_1^*(i)}{\partial \alpha_1} \rho_1^H \mathbf{e}(i) \right] + \text{Re} \left[\tilde{a}_1^*(i) \frac{\partial \rho_1^H}{\partial \alpha_1} \mathbf{e}(i) \right] \right\} \tag{B.6}
\end{aligned}$$

When $l=2$, then

$$\begin{aligned}
\frac{\partial \ln L}{\partial \alpha_2} & = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left(\begin{bmatrix} 0 & \frac{\partial \tilde{a}_2^*(i)}{\partial \alpha_2} & \dots & 0 \end{bmatrix} \begin{bmatrix} \rho_1^p & \rho_1^{p+1} & \dots & \rho_1^{N-1} \\ \rho_2^p & \rho_2^{p+1} & \dots & \rho_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_M^p & \rho_M^{p+1} & \dots & \rho_M^{N-1} \end{bmatrix}^* \mathbf{e}(i) \right) \right\} \\
& + \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left[\tilde{a}_1^*(i) \quad \tilde{a}_2^*(i) \quad \dots \quad \tilde{a}_M^*(i) \right] \begin{bmatrix} 0 & 0 & \dots & 0 \\ \frac{\partial \rho_2^p}{\partial \alpha_2} & \frac{\partial \rho_2^{p+1}}{\partial \alpha_2} & \dots & \frac{\partial \rho_2^{N-1}}{\partial \alpha_2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}^* \mathbf{e}(i) \right\} \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left[\frac{\partial \tilde{a}_2^*(i)}{\partial \alpha_2} \rho_2^H \mathbf{e}(i) \right] + \text{Re} \left[\tilde{a}_2^*(i) \frac{\partial \rho_2^H}{\partial \alpha_2} \mathbf{e}(i) \right] \right\} \tag{B.7}
\end{aligned}$$

Collecting all the above terms for $l=1,2, \dots, M$, yields

$$\begin{aligned}
 \frac{\partial \ln L}{\partial \alpha} &= \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\begin{array}{c} \frac{\partial \tilde{a}_1^*(i)}{\partial \alpha_1} \rho_1^H \\ \frac{\partial \tilde{a}_2^*(i)}{\partial \alpha_2} \rho_2^H \\ \vdots \\ \frac{\partial \tilde{a}_M^*(i)}{\partial \alpha_M} \rho_M^H \end{array} \right] e(i) + \operatorname{Re} \left[\begin{array}{c} \tilde{a}_1^* \frac{\partial \rho_1^H}{\partial \alpha_1} \\ \tilde{a}_2^* \frac{\partial \rho_2^H}{\partial \alpha_2} \\ \vdots \\ \tilde{a}_M^* \frac{\partial \rho_M^H}{\partial \alpha_M} \end{array} \right] e(i) \right\} \\
 &= \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \operatorname{Re} \left[\begin{array}{ccc} \frac{\partial \tilde{a}_1^*(i)}{\partial \alpha_1} & & 0 \\ & \frac{\partial \tilde{a}_2^*(i)}{\partial \alpha_2} & \\ & & \ddots \\ 0 & & \frac{\partial \tilde{a}_M^*(i)}{\partial \alpha_M} \end{array} \right] \begin{bmatrix} \rho_1^H \\ \rho_2^H \\ \vdots \\ \rho_M^H \end{bmatrix} e(i) \\
 &+ \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \operatorname{Re} \left[\begin{array}{ccc} \tilde{a}_1^*(i) & & 0 \\ & \tilde{a}_2^*(i) & \\ & & \ddots \\ 0 & & \tilde{a}_M^*(i) \end{array} \right] \begin{bmatrix} \frac{\partial \rho_1^H}{\partial \alpha_1} \\ \frac{\partial \rho_2^H}{\partial \alpha_2} \\ \vdots \\ \frac{\partial \rho_M^H}{\partial \alpha_M} \end{bmatrix} e(i) \\
 &= \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left(\frac{\partial [\tilde{a}_1^*(i) \ \tilde{a}_2^*(i) \ \dots \ \tilde{a}_M^*(i)]}{\partial \alpha} R^H e(i) \right) + \operatorname{Re} (A_{diag}^H(i) R_\alpha^H e(i)) \right\}
 \end{aligned}$$

$$= \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H \mathbf{e}(i) \right] + \operatorname{Re} \left[A_{diag}^H(i) R_\alpha^H \mathbf{e}(i) \right] \right\} \quad (\text{B.8})$$

7) Proof of (3-16)

From (3-2) we get

$$\begin{aligned} \frac{\partial \ln L}{\partial \omega_l} &= -\frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \left[\frac{\partial}{\partial \omega_l} (\mathbf{y}(i) - R\tilde{\mathbf{a}}(i) + Yd)^H \right] \mathbf{e}(i) + \mathbf{e}^H(i) \left[\frac{\partial}{\partial \omega_l} (\mathbf{y}(i) - R\tilde{\mathbf{a}}(i) + Yd) \right] \right\} \\ &= -\frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \left[-\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega_l} R^H - \tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \omega_l} \right] \mathbf{e}(i) + \mathbf{e}^H(i) \left[-\frac{\partial R}{\partial \omega_l} \tilde{\mathbf{a}}(i) - R \frac{\partial \tilde{\mathbf{a}}(i)}{\partial \omega_l} \right] \right\} \\ &= \frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega_l} R^H \mathbf{e}(i) + \tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \omega_l} \mathbf{e}(i) + \mathbf{e}^H(i) \frac{\partial R}{\partial \omega_l} \tilde{\mathbf{a}}(i) + \mathbf{e}^H(i) R \frac{\partial \tilde{\mathbf{a}}(i)}{\partial \omega_l} \right\} \\ &= \frac{1}{\sigma_e^2} \sum_{i=1}^Q \left\{ \frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega_l} R^H \mathbf{e}(i) + \tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \omega_l} \mathbf{e}(i) + \left(\tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \omega_l} \mathbf{e}(i) \right)^H + \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega_l} R^H \mathbf{e}(i) \right)^H \right\} \\ &= \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega_l} R^H \mathbf{e}(i) \right] + \operatorname{Re} \left[\tilde{\mathbf{a}}^H(i) \frac{\partial R^H}{\partial \omega_l} \mathbf{e}(i) \right] \right\} \quad l = 1, 2, \dots, M \quad (\text{B.9}) \end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \omega_1} &= \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\begin{array}{c} \frac{\partial \tilde{a}_1^*(i)}{\partial \omega_1} \quad 0 \quad \dots \quad 0 \\ \left[\begin{array}{cccc} \rho_1^p & \rho_1^{p+1} & \dots & \rho_1^{N-1} \\ \rho_2^p & \rho_2^{p+1} & \dots & \rho_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_M^p & \rho_M^{p+1} & \dots & \rho_M^{N-1} \end{array} \right]^* \end{array} \right] \mathbf{e}(i) \right\} \\
&+ \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\begin{array}{c} \tilde{a}_1^*(i) \quad \tilde{a}_2^*(i) \quad \dots \quad \tilde{a}_M^*(i) \\ \left[\begin{array}{cccc} \frac{\partial \rho_1^p}{\partial \omega_1} & \frac{\partial \rho_1^{p+1}}{\partial \omega_1} & \dots & \frac{\partial \rho_1^{N-1}}{\partial \omega_1} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right]^* \end{array} \right] \mathbf{e}(i) \right\} \\
&= \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\frac{\partial \tilde{a}_1^*(i)}{\partial \omega_1} \boldsymbol{\rho}_1^H \mathbf{e}(i) \right] + \operatorname{Re} \left[\tilde{a}_1^*(i) \frac{\partial \boldsymbol{\rho}_1^H}{\partial \omega_1} \mathbf{e}(i) \right] \right\} \tag{B.10}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \omega_2} &= \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\begin{array}{c} 0 \quad \frac{\partial \tilde{a}_2^*(i)}{\partial \omega_2} \quad \dots \quad 0 \\ \left[\begin{array}{cccc} \rho_1^p & \rho_1^{p+1} & \dots & \rho_1^{N-1} \\ \rho_2^p & \rho_2^{p+1} & \dots & \rho_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_M^p & \rho_M^{p+1} & \dots & \rho_M^{N-1} \end{array} \right]^* \end{array} \right] \mathbf{e}(i) \right\} \\
&+ \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\begin{array}{c} \tilde{a}_1^*(i) \quad \tilde{a}_2^*(i) \quad \dots \quad \tilde{a}_M^*(i) \\ \left[\begin{array}{cccc} 0 & 0 & \dots & 0 \\ \frac{\partial \rho_2^p}{\partial \omega_2} & \frac{\partial \rho_2^{p+1}}{\partial \omega_2} & \dots & \frac{\partial \rho_2^{N-1}}{\partial \omega_2} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right]^* \end{array} \right] \mathbf{e}(i) \right\}
\end{aligned}$$

$$= \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \text{Re} \left[\frac{\partial \tilde{a}_2^*(i)}{\partial \omega_2} \rho_2^H e(i) \right] + \text{Re} \left[\tilde{a}_2^*(i) \frac{\partial \rho_2^H}{\partial \omega_2} e(i) \right] \right\} \quad (\text{B.11})$$

Collecting all above terms for $l=1,2,\dots,M$, yields

$$\frac{\partial \ln L}{\partial \omega} = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \text{Re} \left(\begin{bmatrix} \frac{\partial \tilde{a}_1^*(i)}{\partial \omega_1} \rho_1^H \\ \frac{\partial \tilde{a}_2^*(i)}{\partial \omega_2} \rho_2^H \\ \vdots \\ \frac{\partial \tilde{a}_M^*(i)}{\partial \omega_M} \rho_M^H \end{bmatrix} e(i) \right) + \text{Re} \left(\begin{bmatrix} \tilde{a}_1^* \frac{\partial \rho_1^H}{\partial \omega_1} \\ \tilde{a}_2^* \frac{\partial \rho_2^H}{\partial \omega_2} \\ \vdots \\ \tilde{a}_M^* \frac{\partial \rho_M^H}{\partial \omega_M} \end{bmatrix} e(i) \right) \right\}$$

$$= \frac{2}{\sigma_e^2} \sum_{i=1}^Q \text{Re} \left(\begin{bmatrix} \frac{\partial \tilde{a}_1^*(i)}{\partial \omega_1} & & & 0 \\ & \frac{\partial \tilde{a}_2^*(i)}{\partial \omega_2} & & \\ & & \ddots & \\ 0 & & & \frac{\partial \tilde{a}_M^*(i)}{\partial \omega_M} \end{bmatrix} \begin{bmatrix} \rho_1^H \\ \rho_2^H \\ \vdots \\ \rho_M^H \end{bmatrix} e(i) \right)$$

$$\begin{aligned}
& + \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \operatorname{Re} \left(\begin{array}{c} \left[\begin{array}{cccc} \tilde{a}_1^*(i) & & & 0 \\ & \tilde{a}_2^*(i) & & \\ & & \ddots & \\ 0 & & & \tilde{a}_M^*(i) \end{array} \right] \begin{bmatrix} \frac{\partial p_1^H}{\partial \omega_1} \\ \frac{\partial p_2^H}{\partial \omega_2} \\ \vdots \\ \frac{\partial p_M^H}{\partial \omega_M} \end{bmatrix} e(i) \end{array} \right) \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left(\frac{\partial [\tilde{a}_1^*(i) \ \tilde{a}_2^*(i) \ \cdots \ \tilde{a}_M^*(i)]}{\partial \omega} R^H e(i) \right) + \operatorname{Re} (A_{diag}^H(i) R_{\omega}^H e(i)) \right\} \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} R^H e(i) \right] + \operatorname{Re} [A_{diag}^H(i) R_{\omega}^H e(i)] \right\} \tag{B.12}
\end{aligned}$$

Appendix C

1) Proof of (3-26a)

From (3-7) and (3-24) we obtain

$$\begin{aligned} E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(i)} \right]^T \right\} &= E \left\{ \frac{4}{\sigma_e^4} \sum_{k=1}^Q \operatorname{Re} [R^H(k) \mathbf{e}(k)] \operatorname{Re} [R^H(i) \mathbf{e}(i)]^T \right\} \\ &= \frac{4}{\sigma_e^4} \frac{1}{2} E \left\{ \operatorname{Re} [R^H \mathbf{e}(k) \mathbf{e}^T(i) R^*] + \operatorname{Re} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\ &= \frac{2}{\sigma_e^4} E \left\{ \operatorname{Re} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\ &= \frac{2}{\sigma_e^2} \operatorname{Re} [R^H R] \delta_{k,i} \end{aligned} \quad (\text{C.1})$$

2) Proof of (3-26b)

From (3-7), (3-8) and (3-24) we obtain

$$\begin{aligned}
 E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_1(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(i)} \right]^T \right\} &= E \left\{ \frac{4}{\sigma_e^4} \sum_{k=1}^Q \text{Im} [R^H(k) \mathbf{e}(k)] \text{Re} [R^H(i) \mathbf{e}(i)]^T \right\} \\
 &= \frac{4}{\sigma_e^4} \frac{1}{2} E \left\{ \text{Im} [R^H \mathbf{e}(k) \mathbf{e}^T(i) R^*] + \text{Im} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\
 &= \frac{2}{\sigma_e^4} E \left\{ \text{Im} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\
 &= \frac{2}{\sigma_e^2} \text{Im} [R^H R] \delta_{k,i} \tag{C.2}
 \end{aligned}$$

3) Proof of (3-27a)

From (3-7), (3-8) and (3-24) we obtain

$$\begin{aligned}
 E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_I(i)} \right]^T \right\} &= E \left\{ \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \operatorname{Re} [R^H(k) \mathbf{e}(k)] \operatorname{Im} [R^H(i) \mathbf{e}(i)]^T \right\} \\
 &= \frac{4}{\sigma_e^4} \frac{1}{2} E \left\{ \operatorname{Im} [R^H \mathbf{e}(k) \mathbf{e}^T(i) R^*] - \operatorname{Im} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\
 &= -\frac{2}{\sigma_e^4} E \left\{ \operatorname{Im} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\
 &= -\frac{2}{\sigma_e^2} \operatorname{Im} [R^H R] \delta_{k,i} \tag{C.3}
 \end{aligned}$$

4) Proof of (3-27b)

From (3-8) and (3-24) we get

$$\begin{aligned}
 E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_1(k)} \right] \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_1(i)} \right]^T \right\} &= E \left\{ \frac{4}{\sigma_e^4} \sum_{k=1}^Q \text{Im} [R^H(k) \mathbf{e}(k)] \text{Im} [R^H(i) \mathbf{e}(i)]^T \right\} \\
 &= -\frac{4}{\sigma_e^4} \frac{1}{2} E \left\{ \text{Re} [R^H \mathbf{e}(k) \mathbf{e}^T(i) R^*] - \text{Re} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\
 &= \frac{2}{\sigma_e^4} E \left\{ \text{Re} [R^H \mathbf{e}(k) \mathbf{e}^H(i) R] \right\} \\
 &= \frac{2}{\sigma_e^2} \text{Re} [R^H R] \delta_{k,i} \tag{C.4}
 \end{aligned}$$

5) Proof of (3-30a)

From (3-5) and (3-24) we obtain

$$\begin{aligned}
 E\left\{-\left[\frac{\partial^2 \ln L}{(\partial \mathbf{d}_R)^2}\right]\right\} &= E\left\{-\frac{\partial}{\partial \mathbf{d}_R}\left(\frac{\partial \ln L}{\partial \mathbf{d}_R}\right)^T\right\} \\
 &= E\left\{-\frac{1}{\sigma_e^2} \sum_{k=1}^Q \frac{\partial}{\partial \mathbf{d}_R} \left[Y^H(k)(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d}) + Y^T(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d})^* \right]^T\right\} \\
 &= E\left\{\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left[(Y^H(k)Y(k))^T + (Y^T(k)Y^*(k))^T \right]\right\} \\
 &= \frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Re}\left[E(Y^H(k)Y(k))\right] \tag{C.5}
 \end{aligned}$$

6) Proof of (3-30b)

From (3-5), (3-6) and (3-24) we get

$$\begin{aligned}
 E\left\{-\left[\frac{\partial^2 \ln L}{\partial \mathbf{d}_I \partial \mathbf{d}_R}\right]\right\} &= E\left\{-\frac{\partial}{\partial \mathbf{d}_I} \left(\frac{\partial \ln L}{\partial \mathbf{d}_R}\right)^T\right\} \\
 &= E\left\{-\frac{1}{\sigma_e^2} \sum_{k=1}^Q \frac{\partial}{\partial \mathbf{d}_I} \left[Y^H(k)(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d}) + Y^T(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d})^*\right]^T\right\} \\
 &= E\left\{\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left[j(Y^H(k)Y(k))^T - j(Y^T(k)Y^*(k))^T\right]\right\} \\
 &= \frac{2}{\sigma_e^2} \sum_{k=1}^Q \text{Im}\left[E(Y^H(k)Y(k))\right] \tag{C.6}
 \end{aligned}$$

7) Proof of (3-31a)

From (3-5), (3-6) and 3-24) we obtain

$$\begin{aligned}
 E\left\{-\left[\frac{\partial^2 \ln L}{\partial \mathbf{d}_R \partial \mathbf{d}_I}\right]\right\} &= E\left\{-\frac{\partial}{\partial \mathbf{d}_R} \left(\frac{\partial \ln L}{\partial \mathbf{d}_I}\right)^T\right\} \\
 &= E\left\{-\frac{1}{\sigma_e^2} \sum_{k=1}^Q \frac{\partial}{\partial \mathbf{d}_R} \left[jY^H(k)(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d}) - jY^T(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d})^* \right]^T\right\} \\
 &= E\left\{-\frac{1}{\sigma_e^2} \sum_{k=1}^Q \left[j(Y^H(k)Y(k))^T - j(Y^T(k)Y^*(k))^T \right]\right\} \\
 &= -\frac{2}{\sigma_e^2} \sum_{k=1}^Q \text{Im} \left[E(Y^H(k)Y(k)) \right] \tag{C.7}
 \end{aligned}$$

8) Proof of (3-31b)

from (3-5) and (3-24) we get

$$\begin{aligned}
 E\left\{-\left[\frac{\partial^2 \ln L}{(\partial \mathbf{d}_1)^2}\right]\right\} &= E\left\{-\frac{\partial}{\partial \mathbf{d}_1}\left(\frac{\partial \ln L}{\partial \mathbf{d}_1}\right)^T\right\} \\
 &= E\left\{-\frac{1}{\sigma_e^2} \sum_{k=1}^Q \frac{\partial}{\partial \mathbf{d}_1} \left[jY^H(k)(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d}) - jY^T(\mathbf{y} - R\tilde{\mathbf{a}} + Y\mathbf{d})^* \right]^T\right\} \\
 &= E\left\{-j \frac{1}{\sigma_e^2} \sum_{k=1}^Q \left[j(Y^H(k)Y(k))^T + j(Y^T(k)Y^*(k))^T \right]\right\} \\
 &= \frac{2}{\sigma_e^2} \sum_{k=1}^Q \operatorname{Re}\left[E(Y^H(k)Y(k))\right] \tag{C.8}
 \end{aligned}$$

9) Proof (3-32a)

From (3-7), (3-17) and (3-24) we get

$$\begin{aligned}
& E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \omega} \right]^T \right\} \\
&= \frac{4}{\sigma_e^4} E \left\{ \text{Re} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^Q \left(\text{Re} \left[A_{diag}^H(i) R_\omega^H \mathbf{e}(i) \right] + \text{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} R^H \mathbf{e}(i) \right] \right) \right]^T \right\} \\
&= \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \text{Re} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^Q \left(\mathbf{e}^T(i) R_\omega^* A_{diag}^*(i) + \mathbf{e}^T(i) R^* \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} \right)^T \right) \right] \right\} \\
&\quad + \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \text{Re} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^Q \left(\mathbf{e}^H(i) R_\omega A_{diag}(i) + \mathbf{e}^H(i) R \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} \right)^H \right) \right] \right\} \\
&= \frac{2}{\sigma_e^2} \left\{ \text{Re} \left[R^H R_\omega A_{diag}(k) \right] + \text{Re} \left[R^H R A_{diag,\omega}(k) \right] \right\} \tag{C.9}
\end{aligned}$$

For the last step we use equation (3-24).

10) Proof of (3-32b)

From (3-8), (3-17) and (3-24) we obtain

$$\begin{aligned}
 & E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_r(k)} \right] \left[\frac{\partial \ln L}{\partial \omega} \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} E \left\{ \text{Im} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^Q \left(\text{Re} \left[A_{diag}^H(i) R_\omega^H \mathbf{e}(i) \right] + \text{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} R^H \mathbf{e}(i) \right] \right) \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \text{Im} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^Q \left(\mathbf{e}^T(i) R_\omega^* A_{diag}^*(i) + \mathbf{e}^T(i) R^* \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} \right)^T \right) \right] \right\} \\
 &\quad + \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \text{Im} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^Q \left(\mathbf{e}^H(i) R_\omega A_{diag}(i) + \mathbf{e}^H(i) R \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} \right)^H \right) \right] \right\} \\
 &= \frac{2}{\sigma_e^2} \left\{ \text{Im} \left[R^H R_\omega A_{diag}(k) \right] + \text{Im} \left[R^H R A_{diag,\omega}(k) \right] \right\} \tag{C.10}
 \end{aligned}$$

11) Proof of (3-33a)

From (3-5), (3-17) and (3-24) we get

$$\begin{aligned}
 E \left\{ - \left[\frac{\partial^2 \ln L}{\partial \mathbf{d}_R(k) \partial \boldsymbol{\omega}} \right] \right\} &= E \left\{ - \frac{\partial}{\partial \mathbf{d}_R(k)} \left(\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right)^T \right\} \\
 &= - \frac{1}{\sigma_s^2} E \left\{ \sum_{i=1}^Q \left[\left(\left[\begin{array}{c} \frac{\partial \tilde{a}_1^*(i)}{\partial \omega_1} \rho_1^H \\ \frac{\partial \tilde{a}_2^*(i)}{\partial \omega_2} \rho_2^H \\ \vdots \\ \frac{\partial \tilde{a}_M^*(i)}{\partial \omega_M} \rho_M^H \end{array} \right] + \left[\begin{array}{c} \tilde{a}_1^*(i) \frac{\partial \rho_1^H}{\partial \omega_1} \\ \tilde{a}_2^*(i) \frac{\partial \rho_2^H}{\partial \omega_2} \\ \vdots \\ \tilde{a}_M^*(i) \frac{\partial \rho_M^H}{\partial \omega_M} \end{array} \right] \right) Y(i) \right]^T \right. \\
 &\quad \left. + Y^H(i) \left(\left[\rho_1 \frac{\partial \tilde{a}_1(i)}{\partial \omega_1} \quad \rho_2 \frac{\partial \tilde{a}_2(i)}{\partial \omega_2} \quad \dots \quad \rho_M \frac{\partial \tilde{a}_M(i)}{\partial \omega_M} \right] + \left[\frac{\partial \rho_1}{\partial \omega_1} \tilde{a}_1(i) \quad \frac{\partial \rho_2}{\partial \omega_2} \tilde{a}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \omega_M} \tilde{a}_M(i) \right] \right) \right\}
 \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^q \left[Y^H(i) \left(\rho_1^* \frac{\tilde{a}_1^H(i)}{\tilde{a}_1} \rho_2^* \frac{\tilde{a}_2^H(i)}{\tilde{a}_2} \dots \rho_M^* \frac{\tilde{a}_M^H(i)}{\tilde{a}_M} \right) + Y^T(i) \left(\frac{\partial \rho_1^*}{\partial \omega_1} \tilde{a}_1^H(i) \frac{\partial \rho_2^*}{\partial \omega_2} \tilde{a}_2^H(i) \dots \frac{\partial \rho_M^*}{\partial \omega_M} \tilde{a}_M^H(i) \right) \right. \right. \\
&\quad \left. \left. + Y^H(i) \left(\left[\rho_1 \frac{\tilde{a}_1(i)}{\tilde{a}_1} \rho_2 \frac{\tilde{a}_2(i)}{\tilde{a}_2} \dots \rho_M \frac{\tilde{a}_M(i)}{\tilde{a}_M} \right] + \left[\frac{\partial \rho_1}{\partial \omega_1} \tilde{a}_1(i) \frac{\partial \rho_2}{\partial \omega_2} \tilde{a}_2(i) \dots \frac{\partial \rho_M}{\partial \omega_M} \tilde{a}_M(i) \right] \right) \right] \right\} \\
&= -\frac{2}{\sigma_e^2} E \left\{ \sum_{i=1}^q \operatorname{Re} \left\{ Y^H(i) \left[\rho_1 \frac{\tilde{a}_1(i)}{\tilde{a}_1} \rho_2 \frac{\tilde{a}_2(i)}{\tilde{a}_2} \dots \rho_M \frac{\tilde{a}_M(i)}{\tilde{a}_M} \right] + Y^H(i) \left[\frac{\partial \rho_1}{\partial \omega_1} \tilde{a}_1(i) \frac{\partial \rho_2}{\partial \omega_2} \tilde{a}_2(i) \dots \frac{\partial \rho_M}{\partial \omega_M} \tilde{a}_M(i) \right] \right\} \right\} \\
&= -\frac{2}{\sigma_e^2} \sum_{i=1}^q \operatorname{Re} \left\{ E(Y^H(i)) R A_{diag, \omega}(i) + E(Y^H(i)) R_\omega A_{diag}(i) \right\} \tag{C.11}
\end{aligned}$$

12) Proof of (3-33b)

From (3-6), (3-17) and (3-24) we get

$$E \left\{ - \left[\frac{\partial^2 \ln L}{\partial \mathbf{d}_I(k) \partial \boldsymbol{\omega}} \right] \right\} = E \left\{ - \frac{\partial}{\partial \mathbf{d}_I(k)} \left(\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right)^T \right\}$$

$$= - \frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^Q j \left[\left(\begin{bmatrix} \frac{\partial \tilde{a}_1^*(i)}{\partial \omega_1} \rho_1^H \\ \frac{\partial \tilde{a}_2^*(i)}{\partial \omega_2} \rho_2^H \\ \vdots \\ \frac{\partial \tilde{a}_M^*(i)}{\partial \omega_M} \rho_M^H \end{bmatrix} + \begin{bmatrix} \tilde{a}_1^*(i) \frac{\partial \rho_1^H}{\partial \omega_1} \\ \tilde{a}_2^*(i) \frac{\partial \rho_2^H}{\partial \omega_2} \\ \vdots \\ \tilde{a}_M^*(i) \frac{\partial \rho_M^H}{\partial \omega_M} \end{bmatrix} \right) Y(i) \right]^T \right\}$$

$$-jY^H(i) \left(\left[\rho_1 \frac{\partial \tilde{a}_1(i)}{\partial \omega_1} \quad \rho_2 \frac{\partial \tilde{a}_2(i)}{\partial \omega_2} \quad \dots \quad \rho_M \frac{\partial \tilde{a}_M(i)}{\partial \omega_M} \right] + \left[\frac{\partial \rho_1}{\partial \omega_1} \tilde{a}_1(i) \quad \frac{\partial \rho_2}{\partial \omega_2} \tilde{a}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \omega_M} \tilde{a}_M(i) \right] \right)$$

$$\begin{aligned}
&= -\frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^Q \left[jY^H(i) \left(\rho_1^* \frac{\partial \tilde{a}_1^H(i)}{\partial \omega_1} \quad \rho_2^* \frac{\partial \tilde{a}_2^H(i)}{\partial \omega_2} \quad \dots \quad \rho_M^* \frac{\partial \tilde{a}_M^H(i)}{\partial \omega_M} \right) + jY^H(i) \left(\frac{\partial \rho_1^*}{\partial \omega_1} \tilde{a}_1^H(i) \quad \frac{\partial \rho_2^*}{\partial \omega_2} \tilde{a}_2^H(i) \quad \dots \quad \frac{\partial \rho_M^*}{\partial \omega_M} \tilde{a}_M^H(i) \right) \right. \right. \\
&\quad \left. \left. - jY^H(i) \left(\rho_1 \frac{\partial \tilde{a}_1(i)}{\partial \omega_1} \quad \rho_2 \frac{\partial \tilde{a}_2(i)}{\partial \omega_2} \quad \dots \quad \rho_M \frac{\partial \tilde{a}_M(i)}{\partial \omega_M} \right) + \left[\frac{\partial \rho_1}{\partial \omega_1} \tilde{a}_1(i) \quad \frac{\partial \rho_2}{\partial \omega_2} \tilde{a}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \omega_M} \tilde{a}_M(i) \right] \right] \right\} \\
&= -\frac{2}{\sigma_e^2} E \left\{ \sum_{i=1}^Q \operatorname{Im} \left\{ Y^H(i) \left[\rho_1 \frac{\partial \tilde{a}_1(i)}{\partial \omega_1} \quad \rho_2 \frac{\partial \tilde{a}_2(i)}{\partial \omega_2} \quad \dots \quad \rho_M \frac{\partial \tilde{a}_M(i)}{\partial \omega_M} \right] + Y^H(i) \left[\frac{\partial \rho_1}{\partial \omega_1} \tilde{a}_1(i) \quad \frac{\partial \rho_2}{\partial \omega_2} \tilde{a}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \omega_M} \tilde{a}_M(i) \right] \right\} \right\} \\
&= -\frac{2}{\sigma_e^2} \sum_{i=1}^Q \operatorname{Im} \left\{ E(Y^H(i)) R A_{diag, \omega}(i) + E(Y^H(i)) R_\omega A_{diag}(i) \right\} \tag{C.12}
\end{aligned}$$

13) Proof of (3-34a)

From (3-7), (3-10) and (3-24) we get

$$\begin{aligned}
 & E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_R(k)} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} E \left\{ \operatorname{Re} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(i) R_{\alpha}^H \mathbf{e}(i) \right] + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\alpha}} R^H \mathbf{e}(i) \right] \right) \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \operatorname{Re} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^{\varrho} \left(\mathbf{e}^T(i) R_{\alpha}^* A_{diag}^*(i) + \mathbf{e}^T(i) R^* \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\alpha}} \right)^T \right) \right] \right\} \\
 &\quad + \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \operatorname{Re} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^{\varrho} \left(\mathbf{e}^H(i) R_{\alpha} A_{diag}(i) + \mathbf{e}^H(i) R \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\alpha}} \right)^H \right) \right] \right\} \\
 &= \frac{2}{\sigma_e^2} \left\{ \operatorname{Re} \left[R^H R_{\alpha} A_{diag}(k) \right] + \operatorname{Re} \left[R^H R A_{diag, \alpha}(k) \right] \right\} \tag{C.13}
 \end{aligned}$$

14) Proof of (3-34b)

From (3-8), (3-10) and (3-24) we obtain

$$\begin{aligned}
 & E \left\{ \left[\frac{\partial \ln L}{\partial \tilde{\mathbf{a}}_1(k)} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} E \left\{ \text{Im} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^{\varrho} \left(\text{Re} \left[A_{diag}^H(i) R_{\alpha}^H \mathbf{e}(i) \right] + \text{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\alpha}} R^H \mathbf{e}(i) \right] \right) \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \text{Im} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^{\varrho} \left(\mathbf{e}^T(i) R_{\alpha}^* A_{diag}^*(i) + \mathbf{e}^T(i) R^* \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\alpha}} \right)^T \right) \right] \right\} \\
 &\quad + \frac{4}{\sigma_e^4} \left(\frac{1}{2} \right) E \left\{ \text{Im} \left[R^H \mathbf{e}(k) \right] \left[\sum_{i=1}^{\varrho} \left(\mathbf{e}^H(i) R_{\alpha} A_{diag}(i) + \mathbf{e}^H(i) R \left(\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\alpha}} \right)^H \right) \right] \right\} \\
 &= \frac{2}{\sigma_e^2} \left\{ \text{Im} \left[R^H R_{\alpha} A_{diag}(k) \right] + \text{Im} \left[R^H R A_{diag, \alpha}(k) \right] \right\} \tag{C.14}
 \end{aligned}$$

15) Proof of (3-35a)

From (3-5), (3-10) and (3-24) we get

$$\begin{aligned}
 E \left\{ - \left[\frac{\partial^2 \ln L}{\partial \mathbf{d}_R(k) \partial \boldsymbol{\alpha}} \right] \right\} &= E \left\{ - \frac{\partial}{\partial \mathbf{d}_R(k)} \left(\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right)^T \right\} \\
 &= - \frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^Q \left(\left[\begin{array}{c} \left[\begin{array}{c} \frac{\partial \tilde{a}_1^*(i)}{\partial \alpha_1} \rho_1^H \\ \frac{\partial \tilde{a}_2^*(i)}{\partial \alpha_2} \rho_2^H \\ \vdots \\ \frac{\partial \tilde{a}_M^*(i)}{\partial \alpha_M} \rho_M^H \end{array} \right] + \left[\begin{array}{c} \tilde{a}_1^*(i) \frac{\partial \rho_1^H}{\partial \alpha_1} \\ \tilde{a}_2^*(i) \frac{\partial \rho_2^H}{\partial \alpha_2} \\ \vdots \\ \tilde{a}_M^*(i) \frac{\partial \rho_M^H}{\partial \alpha_M} \end{array} \right] Y(i) \end{array} \right)^T \right. \\
 &\quad \left. + Y^H(i) \left(\left[\rho_1 \frac{\partial \tilde{a}_1(i)}{\partial \alpha_1} \quad \rho_2 \frac{\partial \tilde{a}_2(i)}{\partial \alpha_2} \quad \dots \quad \rho_M \frac{\partial \tilde{a}_M(i)}{\partial \alpha_M} \right] + \left[\frac{\partial \rho_1}{\partial \alpha_1} \tilde{a}_1(i) \quad \frac{\partial \rho_2}{\partial \alpha_2} \tilde{a}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \alpha_M} \tilde{a}_M(i) \right] \right) \right\}
 \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^Q \left[Y^H(i) \left(\rho_1^* \frac{\tilde{a}_1^H(i)}{\tilde{\alpha}_1} \rho_2^* \frac{\tilde{a}_2^H(i)}{\tilde{\alpha}_2} \dots \rho_M^* \frac{\tilde{a}_M^H(i)}{\tilde{\alpha}_M} \right) + Y^H(i) \left(\frac{\partial \rho_1^*}{\partial \alpha_1} \tilde{a}_1^H(i) \frac{\partial \rho_2^*}{\partial \alpha_2} \tilde{a}_2^H(i) \dots \frac{\partial \rho_M^*}{\partial \alpha_M} \tilde{a}_M^H(i) \right) \right. \right. \\
&\quad \left. \left. + Y^H(i) \left[\rho_1 \frac{\tilde{a}_1(i)}{\tilde{\alpha}_1} \rho_2 \frac{\tilde{a}_2(i)}{\tilde{\alpha}_2} \dots \rho_M \frac{\tilde{a}_M(i)}{\tilde{\alpha}_M} \right] + \left[\frac{\partial \rho_1}{\partial \alpha_1} \tilde{a}_1(i) \frac{\partial \rho_2}{\partial \alpha_2} \tilde{a}_2(i) \dots \frac{\partial \rho_M}{\partial \alpha_M} \tilde{a}_M(i) \right] \right] \right\} \\
&= -\frac{2}{\sigma_e^2} E \left\{ \sum_{i=1}^Q \operatorname{Re} \left\{ Y^H(i) \left[\rho_1 \frac{\tilde{a}_1(i)}{\tilde{\alpha}_1} \rho_2 \frac{\tilde{a}_2(i)}{\tilde{\alpha}_2} \dots \rho_M \frac{\tilde{a}_M(i)}{\tilde{\alpha}_M} \right] + Y^H(i) \left[\frac{\partial \rho_1}{\partial \alpha_1} \tilde{a}_1(i) \frac{\partial \rho_2}{\partial \alpha_2} \tilde{a}_2(i) \dots \frac{\partial \rho_M}{\partial \alpha_M} \tilde{a}_M(i) \right] \right\} \right\} \\
&= -\frac{2}{\sigma_e^2} \sum_{i=1}^Q \operatorname{Re} \left\{ E(Y^H(i)) R A_{diag,\alpha}(i) + E(Y^H(i)) R_\omega A_{diag}(i) \right\} \tag{C.15}
\end{aligned}$$

16) Proof of (3-35b)

From (3-6), (3-10) and (3-24) we obtain

$$\begin{aligned}
 E \left\{ - \left[\frac{\partial^2 \ln L}{\partial \mathbf{d}_1(k) \partial \boldsymbol{\alpha}} \right] \right\} &= E \left\{ - \frac{\partial}{\partial \mathbf{d}_1(k)} \left(\frac{\partial \ln L}{\partial \boldsymbol{\alpha}} \right)^T \right\} \\
 &= - \frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^Q j \left[\left(\left[\begin{array}{c} \frac{\partial \tilde{\mathbf{a}}_1^*(i)}{\partial \alpha_1} \rho_1^H \\ \frac{\partial \tilde{\mathbf{a}}_2^*(i)}{\partial \alpha_2} \rho_2^H \\ \vdots \\ \frac{\partial \tilde{\mathbf{a}}_M^*(i)}{\partial \alpha_M} \rho_M^H \end{array} \right] + \left[\begin{array}{c} \tilde{\mathbf{a}}_1^*(i) \frac{\partial \rho_1^H}{\partial \alpha_1} \\ \tilde{\mathbf{a}}_2^*(i) \frac{\partial \rho_2^H}{\partial \alpha_2} \\ \vdots \\ \tilde{\mathbf{a}}_M^*(i) \frac{\partial \rho_M^H}{\partial \alpha_M} \end{array} \right] \right) Y(i) \right. \\
 &\quad \left. - j Y^H(i) \left(\left[\rho_1 \frac{\partial \tilde{\mathbf{a}}_1(i)}{\partial \alpha_1} \quad \rho_2 \frac{\partial \tilde{\mathbf{a}}_2(i)}{\partial \alpha_2} \quad \dots \quad \rho_M \frac{\partial \tilde{\mathbf{a}}_M(i)}{\partial \alpha_M} \right] + \left[\frac{\partial \rho_1}{\partial \alpha_1} \tilde{\mathbf{a}}_1(i) \quad \frac{\partial \rho_2}{\partial \alpha_2} \tilde{\mathbf{a}}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \alpha_M} \tilde{\mathbf{a}}_M(i) \right] \right) \right\}
 \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{\sigma_e^2} E \left\{ \sum_{i=1}^{\varrho} \left[jY^H(i) \left(\rho_1^* \frac{\tilde{\alpha}_1^H(i)}{\tilde{\alpha}_1} \quad \rho_2^* \frac{\tilde{\alpha}_2^H(i)}{\tilde{\alpha}_2} \quad \dots \quad \rho_M^* \frac{\tilde{\alpha}_M^H(i)}{\tilde{\alpha}_M} \right) + jY^H(i) \left(\frac{\partial \rho_1^*}{\partial \alpha_1} \tilde{\alpha}_1^H(i) \quad \frac{\partial \rho_2^*}{\partial \alpha_2} \tilde{\alpha}_2^H(i) \quad \dots \quad \frac{\partial \rho_M^*}{\partial \alpha_M} \tilde{\alpha}_M^H(i) \right) \right. \right. \\
&\quad \left. \left. - jY^H(i) \left[\rho_1 \frac{\partial \tilde{\alpha}_1(i)}{\partial \alpha_1} \quad \rho_2 \frac{\partial \tilde{\alpha}_2(i)}{\partial \alpha_2} \quad \dots \quad \rho_M \frac{\partial \tilde{\alpha}_M(i)}{\partial \alpha_M} \right] + \left[\frac{\partial \rho_1}{\partial \alpha_1} \tilde{\alpha}_1(i) \quad \frac{\partial \rho_2}{\partial \alpha_2} \tilde{\alpha}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \alpha_M} \tilde{\alpha}_M(i) \right] \right] \right\} \\
&= -\frac{2}{\sigma_e^2} E \left\{ \sum_{i=1}^{\varrho} \operatorname{Im} \left\{ Y^H(i) \left[\rho_1 \frac{\partial \tilde{\alpha}_1(i)}{\partial \alpha_1} \quad \rho_2 \frac{\partial \tilde{\alpha}_2(i)}{\partial \alpha_2} \quad \dots \quad \rho_M \frac{\partial \tilde{\alpha}_M(i)}{\partial \alpha_M} \right] + Y^H(i) \left[\frac{\partial \rho_1}{\partial \alpha_1} \tilde{\alpha}_1(i) \quad \frac{\partial \rho_2}{\partial \alpha_2} \tilde{\alpha}_2(i) \quad \dots \quad \frac{\partial \rho_M}{\partial \alpha_M} \tilde{\alpha}_M(i) \right] \right\} \right\} \\
&= -\frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \operatorname{Im} \left\{ E(Y^H(i)) R A_{\text{diag}, \alpha}(i) + E(Y^H(i)) R_w A_{\text{diag}}(i) \right\} \tag{C.16}
\end{aligned}$$

17) Proof of (3-36)

From (3-17), (3-10) and (3-24) we get

$$\begin{aligned}
 & E \left\{ \left[\frac{\partial \ln L}{\partial \omega} \right] \left[\frac{\partial \ln L}{\partial \alpha} \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} E \left\{ \sum_{k=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(k) R_w^H \mathbf{e}(k) \right] + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \omega} R^H \mathbf{e}(k) \right] \right) \right\} \\
 & \quad \times E \left\{ \sum_{i=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(i) R_a^H \mathbf{e}(i) \right]^T + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H \mathbf{e}(i) \right]^T \right) \right\} \\
 &= \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \omega} R^H \mathbf{e}(k) \right] \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H \mathbf{e}(i) \right]^T \right\} \\
 & \quad + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}(k)}{\partial \omega} R^H \mathbf{e}(k) \right] \operatorname{Re} \left[A_{diag}^H(i) R_a^H \mathbf{e}(i) \right]^T \right\}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[A_{diag}^H(k) R_w^H \mathbf{e}(k) \right] \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H \mathbf{e}(i) \right]^H \right\} \\
& + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[A_{diag}^H(k) R_w^H \mathbf{e}(k) \right] \operatorname{Re} \left[A_{diag}^H(i) R_\alpha^H \mathbf{e}(i) \right]^H \right\} \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[A_{diag,\omega}^*(i) R^H R A_{diag,\alpha}(i) \right] + \operatorname{Re} \left[A_{diag,\omega}^*(i) R^H R_\alpha A_{diag}(i) \right] \right\} \\
& + \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \operatorname{Re} \left[A_{diag}^H(i) R_w^H R A_{diag,\alpha}(i) \right] + \operatorname{Re} \left[A_{diag}^H(i) R_w^H R_\alpha A_{diag}(i) \right] \right\} \quad (\text{C.17})
\end{aligned}$$

18) Proof of (3-37)

From (3-10) and (3-24) we get

$$\begin{aligned}
 & E \left\{ \left[\frac{\partial \ln L}{\partial \alpha} \right] \left[\frac{\partial \ln L}{\partial \alpha} \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} E \left\{ \sum_{k=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(k) R_\alpha^H e(k) \right] + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \alpha} R^H e(k) \right] \right) \right\} \\
 & \quad \times E \left\{ \sum_{i=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(i) R_\alpha^H e(i) \right]^T + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H e(i) \right]^T \right) \right\} \\
 &= \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \alpha} R^H e(k) \right] \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H e(i) \right]^T \right\} \\
 & \quad + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \alpha} R^H e(k) \right] \operatorname{Re} \left[A_{diag}^H(i) R_\alpha^H e(i) \right]^T \right\}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{4}{\sigma_e^4} \sum_{k=1}^Q \sum_{i=1}^Q E \left\{ \text{Re} \left[A_{diag}^H(k) R_\alpha^H \mathbf{e}(k) \right] \text{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \alpha} R^H \mathbf{e}(i) \right]^H \right\} \\
& + \frac{4}{\sigma_e^4} \sum_{k=1}^Q \sum_{i=1}^Q E \left\{ \text{Re} \left[A_{diag}^H(k) R_\alpha^H \mathbf{e}(k) \right] \text{Re} \left[A_{diag}^H(i) R_\alpha^H \mathbf{e}(i) \right]^H \right\} \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \text{Re} \left[A_{diag,\alpha}^*(i) R^H R A_{diag,\alpha}(i) \right] + \text{Re} \left[A_{diag,\alpha}^*(i) R^H R_\alpha A_{diag}(i) \right] \right\} \\
& + \frac{2}{\sigma_e^2} \sum_{i=1}^Q \left\{ \text{Re} \left[A_{diag}^H(i) R_\alpha^H R A_{diag,\alpha}(i) \right] + \text{Re} \left[A_{diag}^H(i) R_\alpha^H R_\alpha A_{diag}(i) \right] \right\} \quad (\text{C.18})
\end{aligned}$$

19) Proof of (3-38)

From (3-17) and (3-24) we get

$$\begin{aligned}
 & E \left\{ \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right] \left[\frac{\partial \ln L}{\partial \boldsymbol{\omega}} \right]^T \right\} \\
 &= \frac{4}{\sigma_e^4} E \left\{ \sum_{k=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(k) R_w^H \mathbf{e}(k) \right] + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \boldsymbol{\omega}} R^H \mathbf{e}(k) \right] \right) \right\} \\
 & \quad \times E \left\{ \sum_{i=1}^{\varrho} \left(\operatorname{Re} \left[A_{diag}^H(i) R_w^H \mathbf{e}(i) \right]^T + \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\omega}} R^H \mathbf{e}(i) \right]^T \right) \right\} \\
 &= \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(k)}{\partial \boldsymbol{\omega}} R^H \mathbf{e}(k) \right] \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \boldsymbol{\omega}} R^H \mathbf{e}(i) \right]^T \right\} \\
 & \quad + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \operatorname{Re} \left[\frac{\partial \tilde{\mathbf{a}}(k)}{\partial \boldsymbol{\omega}} R^H \mathbf{e}(k) \right] \operatorname{Re} \left[A_{diag}^H(i) R_w^H \mathbf{e}(i) \right]^T \right\}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \text{Re} \left[A_{diag}^H(k) R_{\omega}^H \mathbf{e}(k) \right] \text{Re} \left[\frac{\partial \tilde{\mathbf{a}}^H(i)}{\partial \omega} R^H \mathbf{e}(i) \right]^H \right\} \\
& + \frac{4}{\sigma_e^4} \sum_{k=1}^{\varrho} \sum_{i=1}^{\varrho} E \left\{ \text{Re} \left[A_{diag}^H(k) R_{\omega}^H \mathbf{e}(k) \right] \text{Re} \left[A_{diag}^H(i) R^H \mathbf{e}(i) \right]^H \right\} \\
& = \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left[A_{diag,\omega}^*(i) R^H R A_{diag}(i) \right] + \text{Re} \left[A_{diag,\omega}^*(i) R^H R_{\omega} A_{diag}(i) \right] \right\} \\
& + \frac{2}{\sigma_e^2} \sum_{i=1}^{\varrho} \left\{ \text{Re} \left[A_{diag}^H(i) R_{\omega}^H R A_{diag,\omega}(i) \right] + \text{Re} \left[A_{diag}^H(i) R_{\omega}^H R_{\omega} A_{diag}(i) \right] \right\} \quad (\text{C.19})
\end{aligned}$$

Appendix D

Proof of (3-47) through (3-51)

In equation (3-41) the matrix J_{11} can be partitioned as

$$J_{11} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \quad (\text{D.1a})$$

where

$$d_{11} = \begin{bmatrix} \overline{\overline{C}} & -\overline{\overline{C}} & & & \\ \overline{\overline{C}} & \overline{\overline{C}} & & & \\ & & \ddots & & \\ & & & \overline{\overline{C}} & -\overline{\overline{C}} \\ & & & \overline{\overline{C}} & \overline{\overline{C}} \end{bmatrix} \quad (\text{D.1b})$$

$$d_{12} = \begin{bmatrix} \overline{\overline{B(1)}} & -\overline{\overline{B(1)}} \\ \overline{\overline{B(1)}} & \overline{\overline{B(1)}} \\ \vdots & \vdots \\ \overline{\overline{B(Q)}} & -\overline{\overline{B(Q)}} \\ \overline{\overline{B(Q)}} & \overline{\overline{B(Q)}} \end{bmatrix} \quad (\text{D.1c})$$

$$d_{21} = \begin{bmatrix} \overline{B^H(1)} & \overline{-B^H(1)} & \dots & \overline{B^H(Q)} & \overline{-B^H(Q)} \\ \overline{B^H(1)} & \overline{B^H(1)} & \dots & \overline{B^H(Q)} & \overline{B^H(Q)} \end{bmatrix} \quad (\text{D.1d})$$

$$d_{22} = \begin{bmatrix} \overline{F} & \overline{-F} \\ \overline{F} & \overline{F} \end{bmatrix} \quad (\text{D.1e})$$

Then by using the standard block matrix inverse formula, we get

$$J_{11}^{-1} = \begin{bmatrix} (d_{11} - d_{12}d_{22}^{-1}d_{21})^{-1} & -(d_{11} - d_{12}d_{22}^{-1}d_{21})^{-1}d_{12}d_{22}^{-1} \\ -(d_{22} - d_{21}d_{11}^{-1}d_{12})^{-1}d_{21}d_{11}^{-1} & (d_{22} - d_{21}d_{11}^{-1}d_{12})^{-1} \end{bmatrix}$$

$$\triangleq \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (\text{D.2})$$

We define

$$F' = F^{-1} \quad \text{and} \quad C' = C^{-1} \quad (\text{D.3})$$

and obtain

$$d_{11}^{-1} = \begin{bmatrix} \overline{C'} & -\overline{\overline{C'}} & & \\ \overline{\overline{C'}} & \overline{C'} & & \\ & & \ddots & \\ & & & \overline{C'} & -\overline{\overline{C'}} \\ & & & \overline{\overline{C'}} & \overline{C'} \end{bmatrix} \quad (\text{D.4})$$

$$d_{22}^{-1} = \begin{bmatrix} \overline{F'} & -\overline{\overline{F'}} \\ \overline{\overline{F'}} & \overline{F'} \end{bmatrix} \quad (\text{D.5})$$

Now we prove (D.4) and (D.5), because

$$\begin{aligned} d_{22}d_{22}^{-1} &= \begin{bmatrix} \overline{F} & -\overline{\overline{F}} \\ \overline{\overline{F}} & \overline{F} \end{bmatrix} \begin{bmatrix} \overline{F'} & -\overline{\overline{F'}} \\ \overline{\overline{F'}} & \overline{F'} \end{bmatrix} \\ &= \begin{bmatrix} \overline{F} \overline{F'} - \overline{\overline{F}} \overline{\overline{F'}} & -\overline{F} \overline{\overline{F'}} - \overline{\overline{F}} \overline{F'} \\ \overline{\overline{F}} \overline{F'} + \overline{F} \overline{\overline{F'}} & -\overline{\overline{F}} \overline{\overline{F'}} + \overline{F} \overline{F'} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{aligned}$$

The last step in the above holds because

$$\begin{aligned}
 FF' &= (\overline{F} + j\overline{\overline{F}})(\overline{F'} + j\overline{\overline{F'}}) \\
 &= (\overline{F}\overline{F'} - \overline{\overline{F}\overline{\overline{F'}}}) + j(\overline{F}\overline{\overline{F'}} + \overline{\overline{F}\overline{F'}}) = I
 \end{aligned}$$

So that

$$\overline{F}\overline{F'} - \overline{\overline{F}\overline{\overline{F'}}} = I$$

$$\overline{\overline{F}\overline{F'}} + \overline{F}\overline{\overline{F'}} = 0$$

This proves equation (D.5). Equation (D.4) can be proved in a similar way.

By using the matrix inverse lemma, the northwest part of (D.2) can be written as follows.

$$\begin{aligned}
 A_{11} &= (d_{11} - d_{12}d_{22}^{-1}d_{21})^{-1} \\
 &= d_{11}^{-1} + d_{11}^{-1}d_{12}(-d_{21}d_{11}^{-1}d_{12} + d_{22})^{-1}d_{21}d_{11}^{-1}
 \end{aligned} \tag{D.6}$$

In order to evaluate (D.6) we first need to get the following:

$$\begin{aligned}
 d_{11}^{-1}d_{12} &= \begin{bmatrix} \overline{C'B(1)} - \overline{C'B(1)} & -\overline{C'B(1)} - \overline{C'B(1)} \\ \overline{C'B(1)} + \overline{C'B(1)} & -\overline{C'B(1)} + \overline{C'B(1)} \\ \vdots & \vdots \\ \overline{C'B(Q)} - \overline{C'B(Q)} & -\overline{C'B(Q)} - \overline{C'B(Q)} \\ \overline{C'B(Q)} + \overline{C'B(Q)} & -\overline{C'B(Q)} + \overline{C'B(Q)} \end{bmatrix} \\
 &= \begin{bmatrix} \overline{C'B(1)} & -\overline{C'B(1)} \\ \overline{C'B(1)} & \overline{C'B(1)} \\ \vdots & \vdots \\ \overline{C'B(Q)} & -\overline{C'B(Q)} \\ \overline{C'B(Q)} & \overline{C'B(Q)} \end{bmatrix} \tag{D.7}
 \end{aligned}$$

$$A_{22} = (d_{22} - d_{21}d_{11}^{-1}d_{12})^{-1}$$

$$= \left\{ \begin{bmatrix} \overline{F} & -\overline{F} \\ \overline{F} & \overline{F} \end{bmatrix} - \begin{bmatrix} \overline{B^H(1)} & -\overline{B^H(1)} & \dots & \overline{B^H(Q)} & -\overline{B^H(Q)} \\ \overline{B^H(1)} & \overline{B^H(1)} & \dots & \overline{B^H(Q)} & \overline{B^H(Q)} \end{bmatrix} \begin{bmatrix} \overline{C'B(1)} & -\overline{C'B(1)} \\ \overline{C'B(1)} & \overline{C'B(1)} \\ \vdots & \vdots \\ \overline{C'B(Q)} & -\overline{C'B(Q)} \\ \overline{C'B(Q)} & \overline{C'B(Q)} \end{bmatrix} \right\}^{-1}$$

$$\begin{aligned}
&= \left[\begin{array}{cc} \overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)} & -\overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)} \\ \overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)} & \overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)} \end{array} \right]^{-1} \\
&= \left[\begin{array}{cc} \overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)^{-1}} & -\overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)^{-1}} \\ \overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)^{-1}} & \overline{\left(F - \sum_{i=1}^Q B^H(i)C' B(i) \right)^{-1}} \end{array} \right]^{-1} \tag{D.8}
\end{aligned}$$

$$d_{21}d_{11}^{-1} = \left[\begin{array}{cc} \overline{\frac{B^H(1)C'}{B^H(1)C'}} & \overline{\frac{-B^H(1)C'}{B^H(1)C'}} \quad \dots \quad \overline{\frac{B^H(Q)C'}{B^H(Q)C'}} & \overline{\frac{-B^H(Q)C'}{B^H(Q)C'}} \\ \overline{\frac{B^H(1)C'}{B^H(1)C'}} & \overline{\frac{B^H(1)C'}{B^H(1)C'}} \quad \dots \quad \overline{\frac{B^H(Q)C'}{B^H(Q)C'}} & \overline{\frac{B^H(Q)C'}{B^H(Q)C'}} \end{array} \right] \tag{D.9}$$

Equation (D.6) can be obtained by using (D.7), (D.8) and (D.9)

$$A_{11} = \left(d_{11} - d_{12}d_{22}^{-1}d_{21} \right)^{-1}$$

$$= \begin{bmatrix} \overline{C + CB(1)TB^H(1)C} & \overline{-C + CB(1)TB^H(1)C} & \dots & \overline{CB(1)TB^H(Q)C} & \overline{-CB(1)TB^H(Q)C} \\ \overline{C + CB(1)TB^H(1)C} & \overline{C + CB(1)TB^H(1)C} & \dots & \overline{CB(1)TB^H(Q)C} & \overline{CB(1)TB^H(Q)C} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \overline{CB(Q)TB^H(1)C} & \overline{-CB(Q)TB^H(1)C} & \dots & \overline{C + CB(Q)TB^H(Q)C} & \overline{-C + CB(Q)TB^H(Q)C} \\ \overline{CB(Q)TB^H(1)C} & \overline{CB(Q)TB^H(1)C} & \dots & \overline{C + CB(Q)TB^H(Q)C} & \overline{C + CB(Q)TB^H(Q)C} \end{bmatrix} \quad (\text{D.10a})$$

where

$$T = \left(F - \sum_{i=1}^Q B^H(i)C'B(i) \right)^{-1} \quad (\text{D.10b})$$

In order to evaluate A_{12} we first need $d_{12}d_{22}^{-1}$

$$d_{12}d_{22}^{-1} = \begin{bmatrix} \overline{B(1)} & \overline{-B(1)} \\ \overline{B(1)} & \overline{B(1)} \\ \vdots & \vdots \\ \overline{B(Q)} & \overline{-B(Q)} \\ \overline{B(Q)} & \overline{B(Q)} \end{bmatrix} \begin{bmatrix} \overline{F'} & \overline{-F'} \\ \overline{F'} & \overline{F'} \end{bmatrix} \quad (\text{D.11})$$

$$= \begin{bmatrix} \overline{\overline{B(1)F'}} & \overline{\overline{-B(1)F'}} \\ \overline{B(1)F'} & \overline{B(1)F'} \\ \vdots & \vdots \\ \overline{B(Q)F'} & \overline{-B(Q)F'} \\ \overline{B(Q)F'} & \overline{B(Q)F'} \end{bmatrix}$$

From (D.10) and (D.11) we get

$$A_{12} = -(d_{11} - d_{12}d_{22}^{-1}d_{21})^{-1} d_{12}d_{22}^{-1}$$

$$= \begin{bmatrix} \overline{\sum_{k=1}^Q [C' \delta_{k,1} + C' B(1)TB^H(k)C] B(k)F'} & \overline{-\sum_{k=1}^Q [C' \delta_{k,1} + C' B(1)TB^H(k)C] B(k)F'} \\ \overline{\sum_{k=1}^Q [C' \delta_{k,1} + C' B(1)TB^H(k)C] B(k)F'} & \overline{\sum_{k=1}^Q [C' \delta_{k,1} + C' B(1)TB^H(k)C] B(k)F'} \\ \vdots & \vdots \\ \overline{\sum_{k=1}^Q [C' \delta_{k,Q} + C' B(Q)TB^H(k)C] B(k)F'} & \overline{-\sum_{k=1}^Q [C' \delta_{k,Q} + C' B(Q)TB^H(k)C] B(k)F'} \\ \overline{\sum_{k=1}^Q [C' \delta_{k,Q} + C' B(Q)TB^H(k)C] B(k)F'} & \overline{\sum_{k=1}^Q [C' \delta_{k,Q} + C' B(Q)TB^H(k)C] B(k)F'} \end{bmatrix}$$

(D.12)

From (D.8) and (D.9) we get

$$\begin{aligned}
A_{21} &= -\left(d_{22} - d_{21}d_{11}^{-1}d_{12}\right)^{-1}d_{21}d_{11}^{-1} \\
&= -\begin{bmatrix} \overline{TB^H(1)C'} & \overline{-TB^H(1)C'} & \dots & \overline{TB^H(Q)C'} & \overline{-TB^H(Q)C'} \\ \overline{TB^H(1)C'} & \overline{TB^H(1)C'} & \dots & \overline{TB^H(Q)C'} & \overline{TB^H(Q)C'} \end{bmatrix} \tag{D.13}
\end{aligned}$$

Substitution of (D.10), (D.8), (D.12), and (D.13) into (D.2) yields J_{11}^{-1} . Equation (3-47) can be proved by using (3-44), (3-43), (D.2) and (3-42)

$$CRLB^{-1}\left(\begin{bmatrix} \alpha^T & \omega^T \end{bmatrix}^T\right) = J_{22} - J_{21}J_{11}^{-1}J_{12}$$

$$\begin{aligned}
&= \begin{bmatrix} \Delta & X \\ X^H & \Omega \end{bmatrix} \begin{bmatrix} \overline{S^H(1)} & \overline{-S^H(1)} & \dots & \overline{S^H(Q)} & \overline{-S^H(Q)} & \overline{\Gamma^H} & \overline{-\Gamma^H} \\ \overline{W^H(1)} & \overline{-W^H(1)} & \dots & \overline{W^H(Q)} & \overline{-W^H(Q)} & \overline{V^H} & \overline{-V^H} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \overline{S(1)} & \overline{W(1)} \\ \overline{S(1)} & \overline{W(1)} \\ \vdots & \vdots \\ \overline{S(Q)} & \overline{W(Q)} \\ \overline{S(Q)} & \overline{W(Q)} \\ \overline{\Gamma} & \overline{V} \\ \overline{\Gamma} & \overline{V} \end{bmatrix} \\
&\tag{D.14}
\end{aligned}$$

We first evaluate $J_{11}^{-1}J_{12}$

$$J_{11}^{-1} J_{12} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \overline{S(1)} & \overline{W(1)} \\ \overline{S(1)} & \overline{W(1)} \\ \vdots & \vdots \\ \overline{S(Q)} & \overline{W(Q)} \\ \overline{S(Q)} & \overline{W(Q)} \\ \overline{\Gamma} & \overline{V} \\ \overline{\Gamma} & \overline{V} \end{bmatrix}$$

$$= \begin{bmatrix} \overline{\sum_{k=1}^Q [C\delta_{k,1} + CB(1)IB^H(k)C][S(k) - B(k)F\Gamma]} & \overline{\sum_{k=1}^Q [C\delta_{k,1} + CB(1)IB^H(k)C][W(k) - B(k)FV]} \\ \overline{\sum_{k=1}^Q [C\delta_{k,1} + CB(1)IB^H(k)C][S(k) - B(k)F\Gamma]} & \overline{\sum_{k=1}^Q [C\delta_{k,1} + CB(1)IB^H(k)C][W(k) - B(k)FV]} \\ \vdots & \vdots \\ \overline{\sum_{k=1}^Q [C\delta_{k,Q} + CB(Q)IB^H(k)C][S(k) - B(k)F\Gamma]} & \overline{\sum_{k=1}^Q [C\delta_{k,Q} + CB(Q)IB^H(k)C][W(k) - B(k)FV]} \\ \overline{\sum_{k=1}^Q [C\delta_{k,Q} + CB(Q)IB^H(k)C][S(k) - B(k)F\Gamma]} & \overline{\sum_{k=1}^Q [C\delta_{k,Q} + CB(Q)IB^H(k)C][W(k) - B(k)FV]} \\ \overline{\sum_{k=1}^Q [I[\Gamma - B^H(k)C]S(k)]} & \overline{\sum_{k=1}^Q [I[V - B^H(k)C]W(k)]} \\ \overline{\sum_{k=1}^Q [I[\Gamma - B^H(k)C]S(k)]} & \overline{\sum_{k=1}^Q [I[V - B^H(k)C]W(k)]} \end{bmatrix}$$

(D.15)

Second we evaluate

$$J_{21}(J_{11}^{-1}J_{12}) =$$

$$\begin{aligned}
& \left[\operatorname{Re} \sum_{k=1}^{\varrho} \left\{ \Gamma^H T [\Gamma - B^H(k)C'S(k)] + \sum_{l=1}^{\varrho} S^H(l) [C'\delta_{l,k} + C'B(l)TB^H(k)C] \llbracket S(k) - B(k)F'\Gamma \rrbracket \right\} \right. \\
& = \left. \operatorname{Re} \sum_{k=1}^{\varrho} \left\{ V^H T [\Gamma - B^H(k)C'S(k)] + \sum_{l=1}^{\varrho} W^H(l) [C'\delta_{l,k} + C'B(l)TB^H(k)C] \llbracket S(k) - B(k)F'\Gamma \rrbracket \right\} \right. \\
& \left. \operatorname{Re} \sum_{k=1}^{\varrho} \left\{ \Gamma^H T [V - B^H(k)C'W(k)] + \sum_{l=1}^{\varrho} S^H(l) [C'\delta_{l,k} + C'B(l)TB^H(k)C] \llbracket W(k) - B(k)F'V \rrbracket \right\} \right] \\
& \left. \operatorname{Re} \sum_{k=1}^{\varrho} \left\{ V^H T [V - B^H(k)C'W(k)] + \sum_{l=1}^{\varrho} W^H(l) [C'\delta_{l,k} + C'B(l)TB^H(k)C] \llbracket W(k) - B(k)F'V \rrbracket \right\} \right] \\
& \tag{D.16}
\end{aligned}$$

We note that (D.16) is a (2×2) block matrix. Finally, by using (D.14), (D.15) and (D.16) we can get (3-47) as follows:

$$\text{CRLB}^{-1} \left(\begin{bmatrix} \alpha^T & \omega^T \end{bmatrix}^T \right) = \begin{bmatrix} \Lambda & X \\ X^H & \Omega \end{bmatrix} - \begin{bmatrix} \Lambda & \Pi \\ \Xi & \Theta \end{bmatrix} \tag{D.17}$$

The $\Lambda, \Xi, \Pi, \Theta$ have been defined in (D.15) and (D.16), and therefore we proved equations (3-48), (3-49), (3-50) and (3-51).

Appendix E

Proof of (4-9) and (4-15)

1) Proof of equation (4-9)

From (3-51) for this special case we get

$$\Theta = \text{Re} \left\{ V^H (F - B^H C^{-1} B)^{-1} (V - B^H C^{-1} W) \right. \\ \left. + W^H \left[C^{-1} + C^{-1} B (F - B^H C^{-1} B)^{-1} B^H C^{-1} \right] (W - B F^{-1} V) \right\} \quad (\text{E.1})$$

Substitution of equations (3-33), (3-30), (3-28), (3-26), and (3-32) into (E.1) yields

$$\Theta = \frac{2}{\sigma_e^2} \text{Re} \left\{ (A_{diag,\omega}^H R^H E(Y) + A_{diag,\omega}^H R_\omega^H E(Y)) T \times \right. \\ \left. \times \left[E(Y^H) R A_{diag,\omega} + E(Y^H) R_\omega A_{diag} - E(Y^H) R (R^H R)^{-1} (R^H R A_{diag,\omega} + R^H R_\omega A_{diag}) \right] \right\}$$

$$\begin{aligned}
& + (A_{diag,\omega}^H R^H R + A_{diag}^H R_{\omega}^H R) \left[(R^H R)^{-1} + (R^H R)^{-1} R^H E(Y) T E(Y^H) R (R^H R)^{-1} \right] \\
& \times \left[(R^H R A_{diag,\omega} + R^H R_{\omega} A_{diag}) - R^H E(Y) (E(Y^H Y))^{-1} (E(Y^H) R A_{diag,\omega} + E(Y^H) R_{\omega} A_{diag}) \right] \Big\} \\
\end{aligned} \tag{E.2}$$

where

$$T = \left[E(Y^H Y) - E(Y^H) R (R^H R)^{-1} R^H E(Y) \right]^{-1} \tag{E.3}$$

Substitution of (4-7) into (E.2) yields

$$\begin{aligned}
& \ominus \frac{2}{\sigma_e^2} \text{Re} \left\{ (-j\tilde{\alpha}^* R^H H(Y) + \tilde{\alpha}^* R_{\omega}^H H(Y)) T \times \right. \\
& \times \left[E(Y^H) R j\tilde{\alpha} + E(Y^H) R_{\omega} \tilde{\alpha} - E(Y^H) R \frac{1}{N-p} \left((N-p)j\tilde{\alpha} + j \frac{(N+p-1)(N-p)}{2} \tilde{\alpha} \right) \right] \\
& \left. + \left[-j\tilde{\alpha}^* (N-p) + \tilde{\alpha}^* \left(-j \frac{(N+p-1)(N-p)}{2} \right) \right] \left[\frac{1}{N-p} + \frac{1}{N-p} R^H E(Y) T E(Y^H) R \frac{1}{N-p} \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& \times \left[(N-p)j\tilde{a} + j \frac{(N+p-1)(N-p)}{2} \tilde{a} - R^H E(Y) (E(Y^H Y))^{-1} (E(Y^H) R j \tilde{a} + E(Y^H) R_{\omega} \tilde{a}) \right] \\
& = \frac{2}{\sigma_e^2} \operatorname{Re} \left\{ \tilde{a}^* \tilde{a} \left(-j R^H E(Y) + R_{\omega}^H E(Y) \right) T \left[E(Y^H) R_{\omega} - j \frac{N+p-1}{2} E(Y^H) R \right] \right. \\
& \quad \left. + \tilde{a}^* \tilde{a} \left[-j \frac{(N-p)(N+p+1)}{2} \frac{1}{N-p} \left(1 + \frac{1}{N-p} R^H E(Y) T E(Y^H) R \right) \right] \right. \\
& \quad \left. \times \left[j \frac{(N-p)(N+p+1)}{2} - j R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R - R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R_{\omega} \right] \right\} \\
& \tag{E.4}
\end{aligned}$$

Rearranging (E.4) yields

$$\begin{aligned}
\Theta & = \frac{2|\tilde{a}|^2}{\sigma_e^2} \operatorname{Re} \left\{ \frac{(N-p)(N+p+1)^2}{4} + \frac{(N+p+1)^2 - 2(N+p-1)}{4} R^H E(Y) T E(Y^H) R \right. \\
& \quad \left. + R_{\omega}^H E(Y) T E(Y^H) R_{\omega} - \frac{(N+p+1)}{2} R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R \right. \\
& \quad \left. - \frac{(N+p+1)}{2(N-p)} R^H E(Y) T E(Y^H) R R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R \right\}
\end{aligned}$$

$$\begin{aligned}
& +j \left[\frac{(N+p+1)}{2} R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R_{\omega} \right. \\
& + \frac{(N+p+1)}{2(N-p)} R^H E(Y) T E(Y^H) R R^H E(Y) (E(Y^H Y))^{-1} E(Y^H) R_{\omega} \\
& \left. - \frac{(N+p-1)}{2} R_{\omega}^H E(Y) T E(Y^H) R - R^H E(Y) T E(Y^H) R_{\omega} \right] \quad (E.5)
\end{aligned}$$

Equation (E.5) is (4-9). Therefore, equation (4-9) is proved.

2) Proof of equation (4-15)

Substitution of (4-14) into (4-11) yields

$$\begin{aligned}
 CRLB(\omega) &= \frac{\sigma_e^2}{|\tilde{a}|^2} \frac{6}{N(N-1)(N-2)} \operatorname{Re} \left\{ 1 - \frac{3}{N(N-1)(N-2)} \right. \\
 &\times \left[\left((N+1)^2 + 3 \right) \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} T \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} + 4 \left(-j \sum_{i=1}^{N-1} i E(y_i) e^{-ji\omega} \right) T j \sum_{k=1}^{N-1} k E(y_k^*) e^{jk\omega} \right. \\
 &- 2(N+2) \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} (T) \sum_{i=1}^{N-1} E(y_i^*) e^{ji\omega} - \frac{2(N+2)}{N-1} \left(\sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} \right)^2 T \left(\sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \right)^2 (T) \\
 &+ j \left(2(N+2) \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} (T) j \sum_{i=1}^{N-1} i E(y_i^*) e^{ji\omega} - 2N \left(-j \sum_{i=1}^{N-1} i E(y_i) e^{-ji\omega} \right) T \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \right. \\
 &\left. \left. \left. + \frac{2(N+2)}{N-1} \left(\sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} \right)^2 T \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} (T) j \sum_{i=1}^{N-1} i E(y_i^*) e^{ji\omega} - 4 \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} T j \sum_{k=1}^{N-1} k E(y_k^*) e^{jk\omega} \right) \right] \right\}^{-1}
 \end{aligned} \tag{E.6a}$$

where

$$T1 = \left(E \left(\sum_{k=1}^{N-1} |y_k|^2 \right) \right)^{-1} \quad (\text{E.6b})$$

From (E.6a) we get

$$\begin{aligned}
CRLB(\omega) &= \frac{\sigma_e^2}{|\tilde{a}|^2} \frac{6}{N(N-1)(N-2)} \operatorname{Re} \left\{ 1 - \frac{3T}{N(N-1)(N-2)} \right. \\
&\times \left[\left((N+2)^2 - 2N \right) \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} + 4 \sum_{i=1}^{N-1} i E(y_i) e^{-ji\omega} \sum_{k=1}^{N-1} k E(y_k^*) e^{jk\omega} \right. \\
&- 2(N+2) T^{-1} \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} (T1) \sum_{i=1}^{N-1} E(y_i^*) e^{ji\omega} - \frac{2(N+2)}{N-1} \left(\sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} \right)^2 \left(\sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \right)^2 (T1) \\
&- 2(N+2) T^{-1} \sum_{i=1}^{N-1} E(y_i) e^{-ji\omega} (T1) \sum_{i=1}^{N-1} i E(y_i^*) e^{ji\omega} - 2N \sum_{i=1}^{N-1} i E(y_i) e^{-ji\omega} \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \\
&\left. \left. - \frac{2(N+2)}{N-1} \left(\sum_{i=1}^{N-1} y_i e^{-ji\omega} \right)^2 \sum_{k=1}^{N-1} y_k^* e^{jk\omega} (T1) \sum_{i=1}^{N-1} t y_i^* e^{jt\omega} + 4 \sum_{i=1}^{N-1} y_i e^{-ji\omega} \sum_{k=1}^{N-1} k y_k^* e^{jk\omega} \right) \right] \right\}^{-1} \quad (\text{E.7})
\end{aligned}$$

Rearrange (E.7) to yield

$$\begin{aligned}
 CRLB(\omega) &= \frac{\sigma_e^2}{|\tilde{a}|^2} \frac{6}{N(N-1)(N-2)} \operatorname{Re} \left\{ 1 - \frac{3T}{N(N-1)(N-2)} \right. \\
 &\quad \times \left[4 \sum_{i=1}^{N-1} E(y_i^*) e^{j\omega} \sum_{k=1}^{N-1} (k+1) E(y_k) e^{-jk\omega} - 2N \sum_{i=1}^{N-1} E(y_i^*) e^{j\omega} \sum_{k=1}^{N-1} (k+1) E(y_k) e^{-jk\omega} \right. \\
 &\quad \left. \left. + (N+2)^2 \sum_{i=1}^{N-1} E(y_i) e^{-j\omega} \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \right. \right. \\
 &\quad \left. \left. \frac{2N+4}{N-1} \sum_{k=1}^{N-1} E(y_k) e^{-jk\omega} (T) \sum_{l=1}^{N-1} (l+1) E(y_l^*) e^{jl\omega} \left((N-1)T^{-1} + \sum_{i=1}^{N-1} E(y_i) e^{-j\omega} \sum_{k=1}^{N-1} E(y_k^*) e^{jk\omega} \right) \right] \right\}^{-1}
 \end{aligned}
 \tag{E.8}$$

Equation (E.8) is (4-15). Therefore, equation (4-15) is proved.

References

- [1] D. W. Tufts and R. Kumaresan, "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood," the IEEE Proceedings, Vol. 70, No. 9, pp. 975-989, Sept. 1982.
- [2] M. Kaveh and A. J. Barabell, "The Statistical Performance of The MUSIC and The Minimum-Norm Algorithms in Resolving Plane Waves in Noise," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, pp. 331-341, Apr. 1986.
- [3] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," in Proc. RADC Spectral Estimation Workshop, (Rome, NY), pp. 243-258, 1986.
- [4] R. Kumaresan, L. L. Scharf, and A. K. Shaw, "An Algorithm for Pole-Zero Modeling and Analysis," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, pp. 637-640, June 1986.
- [5] Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, pp. 1361-1375, Oct. 1986.

- [6] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 37, pp. 984-995, July 1989.
- [7] R. Kumaresan, D. W. Tufts, and Louis L. Scharf, "A Prony Method for Noise Data: Choosing the Signal Components and Selecting the Order in Exponential Signal Models," *Proc. IEEE*, Vol. 72, No. 2, pp. 230-233, Feb. 1984.
- [8] T. J. Abatzoglou and L. K. Lam, "Direction Finding Using Uniform Arrays and the Constrained Total Least Squares Method," in *Proc. of the twenty-fifth Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 4-6, 1991.
- [9] S. W. Lang and J. H. McClellan, "Frequency Estimation with Maximum Entropy Spectral Estimations," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-28, pp. 716-724, Dec. 1980.
- [10] F. Tuteur and Y. Rockah, "The Covariance Difference Method in Signal Detection," in *Proc. 3rd ASSP Workshop on Spectrum Estimation and Modeling* Boston, MA, , pp. 120-122, Nov. 1986.
- [11] P. Swerling, "Probability of Detection of Fluctuating Target," *IRE Trans. Inform. Theory*, 6, pp. 269-308, 1960.

- [12] Q. T. Zhang and Simon Haykin, "Model-Based Detection," Adaptive Radar Detection and Estimation, Edited by Simon Haykin and Allan Steinhardt. John Wiley & Sons, Inc. 1992.
- [13] H. L. Van Trees, Detection, Estimation and Modulation Theory, part 1, New York; Wiley, 1968.
- [14] C. R. Rao, "Minimum Variance and the Estimation of Several Parameters," Proc. Cambridge Phil. Soc., Vol. 43, pp. 280-283, 1946.
- [15] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramer-Rao bound," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-37, pp. 720-741, May 1989.
- [16] William M. Steedly and Randolph L. Moses, "The Cramer-Rao Bound for Pole and Amplitude Coefficient Estimates of Damped Exponential Signals in Noise," IEEE Trans. Acoustics, Speech, Signal Processing, Vol. ASSP-41, No. 3, pp. 1305-1318, March 1993.
- [17] D. C. Rife and R. R. Boorstyn, "Signal-tone Parameter Estimation from Discrete-time Observations," IEEE Trans. Inform. Theory, Vol. IT-20, pp. 591-598, Sept. 1974.

- [18] D. C. Rife and R. R. Boorstyn, "Multiple Tone Parameter Estimation from Discrete-time Observations," *Bell Syst. Tech. J.*, Vol. 55, pp. 1389-1410, Nov. 1976.
- [19] Tong-Zhang Kou and A. A. (Louis) Beex, "Parameter Identification of Exponential Signals in Colored Noise Environments," 8-th IFAC/IFORS Symposium on Identification and System Parameter Estimation, Beijing PRC, pp. 1415-1420, 27-31 Aug. 1988.
- [20] S. M. Kay, Modern Spectral Estimation, New Jersey: Prentice Hall, 1988.
- [21] Athanasios Papoulis, Probability, Random Variables, and Stochastic Processes, Second Edition, McGraw-Hill, Inc. 1984.
- [22] G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control, Holdenday, San Francisco, 1970.
- [23] W. M. Steedly and R. L. Moses, "High resolution exponential modeling of fully polarized radar returns," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 27, pp. 459-469, May 1991.
- [24] T. McWhorter and L. L. Scharf, "Cramer-Rao Bound for Deterministic Modal Analysis," *IEEE Trans. on Signal Processing*, Vol. 41, No 5, pp. 1847-1866, May 1993.

- [25] S. J. Rabinowitz, C. H. Gager, E. Brookner, C. E. Muehe, and C. M. Johnson, "Applications of Digital Technology to Radar," Proc. IEEE 73:2, pp. 325-339, Feb. 1985.
- [26] J. W. Taylor and G. Brunins, "Design of a New Airport Surveillance Radar (ASR-9)," Proc. IEEE 73.2, pp. 284-289, Feb. 1985.
- [27] Chang, C. B. and Tabaczynski, "Application of State Estimation to Target Tracking," IEEE Trans. on Automatic Control, Vol. 29, No. 2, 4, pp. 98-109, Feb. 1984.
- [28] W. C. Knight, R. C. Pridham, and S. M. Kouy, "Digital Signal Processing for Sonar," Proc. IEEE 96, pp. 1451-1506, 1981.
- [29] A. B. Baggeroer, "Sonar Signal Processing," in Applications of Digital Signal Processing, ed. A. V. Oppenheim, Englewood Cliffs, N. J.; Prentice-Hall 1978.
- [30] Y. T. Chan, ed., Underwater Acoustic Data Processing, Dordrecht, the Netherlands: Kluwer, 1988.
- [31] A. M. Bruckstein, T. J. Shan, and T. Kailath, "The Resolution of Overlapping Echos," IEEE Trans. Acoustics, Speech, Signal Processing, Vol. Assp-33, pp. 1357-1367, Dec. 1985.

- [32] S. Haykin, ed., Array Signal Processing, Englewood Cliffs, N. J.; Prentice-Hall, 1985.
- [33] G. Biennu and C. Kopp, "Adaptivity to Background Noise Spatial Coherence for High Resolution Passive Methods," in Proc. IEEE ICASSP Denver, Co, pp. 307-310, 1990.
- [34] R. O. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation," Ph. D. dissertation, Stanford Univ., Stanford, CA, Nov. 1981.
- [35] B. Porat and B. Friedlander, "Analysis of the Asymptotic Relative Efficiency of the MUSIC Algorithm," IEEE Trans. Acoustics, Speech, Signal Processing, Vol. ASSP-36, pp. 532-544, Apr. 1988.
- [36] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT-A Subspace Rotation Approach to Estimation of Parameters of Cisoids in Noise." IEEE Trans. Acoustics, Speech, Signal Processing, Vol. ASSP-34, No. 4, pp. 1340-1342, Oct. 1986.
- [37] R. Ray and T. Kailath, "Total Least Squares ESPRIT," Proc. 21st Asilomar Conf. on Circuits, Systems and Computers, Pacific Grove, CA, pp. 297-301, Nov. 1987.
- [38] Y. Hua and T. K. Sarkar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped/Undamped Sinusoids in Noise," IEEE Trans. Acoustics, Speech, Signal Processing, Vol. 38, pp. 814-824, May 1990.

- [39] G. F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function," *Geophysics Journal Royal Astronomical Soc.* 33, pp. 347-366, 1973.
- [40] H. Akaike, "Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models," *Biometrika* 60, pp. 255-265, 1973.
- [41] T. W. Anderson, "Statistical Inference for Covariance Matrices with Linear Structure," *Multivariate Analysis-II*, P. R. Krishnaiah, Ed. New York: Academic Press, 1969.
- [42] C. J. Gueguen and L. L. Scharf, "Exact Maximum Likelihood Identification of ARMA Models: A Signal Processing Perspective," *Proc. EUSIPCO*, pp. 759-769, Sept. 1980.
- [43] S. Kay, "Recursive Maximum Likelihood Estimation of Autoregressive Processes," *IEEE Trans. ASSP-31*, pp. 56-65, Feb. 1983.
- [44] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of Structured Covariance Matrices," *IEEE Proc.* 70, pp. 963-974, Sept. 1982.
- [45] D. R. Fuhrmann and M. I. Miller, "On the Existence of Positive Definite Maximum Likelihood Estimates of Structured Covariance Matrices," *IEEE Trans. IT-34*, pp. 722-729, July 1988.

- [46] J. P. Le Cadre, "Parametric Method for Spatial Signal Processing in unknown Colored Noise Fields," IEEE Trans. Vol. 37, No. 7, pp. 965-983, July 1989.
- [47] J. Bohm and D. Kraus, "On Least Squares Methods for DOA Estimation in Unknown Noise Field," Proc. ICASSP, pp. 2833-2836, 1988.
- [48] Q. Wu et al, "Maximum Likelihood Estimation for Array Processing in Unknown Noise Environments," Proc. 1992 ICASSP, pp. V. 241-V. 244, 1992.
- [49] M. Wax, "Detection and Localization of Multiple Sources in Noise with Unknown Covariance," IEEE Trans. Signal Processing, Vol. 40, No. 1, pp. 245-249, Jan. 1992.
- [50] C. Chatterjee, R. L. Kashyap and G. K. Boray, "Estimation of Close Sinusoids in Colored Noise and Model Discrimination," IEEE Trans. Acoustics, Speech, Signal Processing, pp. 328-337, March 1987.
- [51] A. A. (Louis) Beex and Louis L. Scharf, "Covariance Sequence Approximation for Parametric Spectrum Modeling," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-29, No. 5, pp. 1042-1052, Oct. 1981.
- [52] F. B. Hildebrand, Introduction to Numerical Analysis, New York: McGraw-Hill, 1956.

- [53] P. Eykhoff, System Identification-parameter and state estimation, John Wiley and Sons, 1974.
- [54] G. C. Goodwin and K. S. Sin, Adaptive Filtering, Prediction, and Control, Prentice-Hall, 1984.
- [55] L. Ljung and T. Söderström, Theory and Practice of Recursive Identification, MIT Press, 1983.
- [56] T. Kailath, Linear System, Prentice-Hall, 1980.
- [57] B. Noble and J. W. Daniel, Applied Linear Algebra, Prentice-Hall, Inc., Englewood Cliffs, N. J. 2nd ed., 1977.
- [58] J. P. Dugre, A. A. (Louis) Beex, and L. L. Sharf, "Generating Covariance Sequences and the Calculation of Quantization and Rounding Error Variances in Digital Filters," IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-28, No. 1, pp. 102-104, Feb. 1980.

VITA

Tong-Zhang Kou was born in Hopei Province, China. He received the B.S. degree from Northwest Telecommunication University, Xian, China, in 1970. From 1970 to 1978 he was employed as an electrical engineer in National Xinghuo Factory, Xinjian, China. In 1982 he received the M.S. degree from Zhejiang University, Hangzhou, China, and became employed as a research engineer in the Changcheng Institute of Metrology and Measuring Technology (CIMM), Beijing, China.

In 1989 he came to the Department of Electrical Engineering, Virginia Polytechnic Institute and State University as a visiting scholar. In 1993 he received the M.S. degree from the Department of Electrical Engineering at Virginia Polytechnic Institute and State University. He is actively involved in signal processing research in the areas of parameter identification, parameter estimation, detection, and model analysis in stochastic signal processing.

A handwritten signature in black ink, reading "Kou Tongzhang". The signature is written in a cursive, flowing style with long, sweeping lines.