Necessary conditions for a maximum likelihood estimate to become asymptotically unbiased and attain the Cramer–Rao Lower Bound. Part I. General approach with an application to time-delay and Doppler shift estimation

Eran Naftali and Nicholas C. Makris
Massachusetts Institute of Technology, Room 5-222, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

(Received 17 October 2000; revised 20 April 2001; accepted 29 May 2001)

Analytic expressions for the first order bias and second order covariance of a general maximum likelihood estimate (MLE) are presented. These expressions are used to determine general analytic conditions on sample size, or signal-to-noise ratio (SNR), that are *necessary* for a MLE to become asymptotically unbiased and attain minimum variance as expressed by the Cramer–Rao lower bound (CRLB). The expressions are then evaluated for multivariate Gaussian data. The results can be used to determine asymptotic biases, variances, and conditions for estimator optimality in a wide range of inverse problems encountered in ocean acoustics and many other disciplines. The results are then applied to rigorously determine conditions on SNR necessary for the MLE to become unbiased and attain minimum variance in the classical active sonar and radar time-delay and Doppler-shift estimation problems. The time-delay MLE is the time lag at the peak value of a matched filter output. It is shown that the matched filter estimate attains the CRLB for the signal's position when the SNR is much larger than the *kurtosis* of the expected signal's energy spectrum. The Doppler-shift MLE exhibits dual behavior for narrow band analytic signals. In a companion paper, the general theory presented here is applied to the problem of estimating the range and depth of an acoustic source submerged in an ocean waveguide.
[DOI: 10.1121/1.1387091]

PACS numbers: 43.30.Pc, 43.60.Gk, 43.60.Pt [DLB]

I. INTRODUCTION

In many practical problems in ocean acoustics, geophysics, statistical signal processing, and other disciplines, nonlinear inversions are required to estimate parameters from measured data that undergo random fluctuations. The nonlinear inversion of random data often leads to estimates that are biased and do not attain minimum variance, namely the Cramer–Rao lower bound (CRLB), for small sample sizes or equivalently low signal-to-noise ratio (SNR). The maximum likelihood estimator (MLE) is widely used because if an asymptotically unbiased and minimum variance estimator exists for large sample sizes, it is guaranteed to be the MLE.[1] Since exact expressions for the bias, variance, and error correlation of the MLE are often difficult or impractical to derive analytically, it has become popular in ocean acoustics and many other areas to simply neglect potential biases and to compute limiting bounds on the mean square error, such as the CRLB, since these bounds are usually much easier to obtain. The CRLB, however, typically provides an unrealistically optimistic approximation to the MLE error correlation in many nonlinear inverse problems when the sample size is small, or equivalently the SNR is low. A number of bounds on the error correlation exist that are tighter than the CRLB.[1–5] Some of these bounds are based on Bayesian assumptions[4,5] and so require the *a priori* probability density of the parameters to be estimated, which can be problematic when the *a priori* probability density is not known.[6]

The purpose of the present paper is not to derive a new parameter resolution bound, but rather to determine, within the framework of classical estimation theory[1,6,7] the conditions on sample size, or SNR, *necessary* for the MLE to become asymptotically unbiased and attain minimum variance. The approach is to apply the tools of higher order asymptotic inference, which rely heavily on tensor analysis, to expand the MLE as a series in inverse orders of sample size or equivalently inverse orders of SNR.[7] From this series analytic expressions for the first order bias, second order covariance and second order error correlation of a general MLE are presented in terms of joint moments of the log-likelihood function and its derivatives with respect to the parameters to be estimated. Since the first order error correlation is shown to be the CRLB, which is only valid for unbiased estimates, the second order error correlation can provide a tighter error approximation to the MLE than the CRLB that is applicable in relatively low SNR even when the MLE is biased to first order. These expressions are then used to determine general analytic requirements on sample size, or SNR, that are *necessary* for an MLE to become asymptotically unbiased and attain minimum variance. This is done by showing when the first order bias becomes negligible compared to the true value of the parameter and when the second order covariance term becomes negligible compared to the CRLB.

The first order bias is evaluated for general multivariate Gaussian data. The second order covariance and error correlation terms are evaluated for two special cases of Gaussian data that are of great practical value in ocean acoustics, geophysics, and statistical signal processing. The first is for a deterministic signal vector embedded in additive noise and the second is for a fully randomized signal vector with zero mean in additive noise. These cases have been widely used in ocean acoustic inversions, spectral estimation, beamforming, sonar and radar detection, and localization problems, as well as statistical optics.[8–10] In a companion paper, each of these cases is applied to determine the asymptotic bias and covariance of maximum likelihood range and depth estimates of a sound source submerged in an ocean waveguide from measured hydrophone array data as well as necessary conditions for the estimates to attain the CRLB.[11]

In the present paper, these expressions are applied to the active sonar and radar time-delay and Doppler-shift estimation problems, where time delay is used for target range estimation and Doppler shift is used for target velocity estimation. Attention is focused on the commonly encountered scenario of a deterministic signal with unknown spatial or temporal delay received together with additive white noise. The time-delay MLE is then the time lag at the peak value of a matched filter output. The matched filter estimate for a signal's time delay or position is widely used in many applications of statistical pattern recognition in sonar, radar, and optical image processing. This is because it has long been known that the matched filter estimate attains the CRLB in high SNR. *Necessary* analytic conditions on how high the SNR must be for the matched filter estimate to attain the CRLB have not been previously obtained but are derived here using the general asymptotic approach developed in Secs. II–IV.

A number of authors have derived tighter bounds than the CRLB for the time-delay estimation problem to help evaluate performance at low SNR where the CRLB is not attained by the MLE, as, for example, in Refs. 5, 12, 13. The present paper follows a different approach by providing explicit expressions for the second order variance of the time-delay and Doppler-shift MLEs that are attained in lower SNR than the CRLB. The first order bias is also derived. These expressions are then used to provide analytic conditions on SNR necessary for the time-delay MLE, namely the matched filter estimate, and Doppler-shift MLE to become unbiased and attain minimum variance in terms of properties of the signal and its spectrum. Illustrative examples for standard linear frequency modulated (LFM), hyperbolic frequency modulated (HFM), and canonical waveforms are provided for typical low-frequency active-sonar scenarios in ocean acoustics.

II. GENERAL ASYMPTOTIC EXPANSIONS OF THE MLE AND ITS MOMENTS

Suppose the random data vector $\mathbf{X}$, given m-dimensional parameter vector $\boldsymbol{\theta}$, obeys the conditional probability density function (pdf) $p(\mathbf{X};\boldsymbol{\theta})$. The log-likelihood function $l(\boldsymbol{\theta})$ is defined as $l(\boldsymbol{\theta})=\ln(p(\mathbf{X};\boldsymbol{\theta}))$ when evaluated at measured values of $\mathbf{X}$. Let the rth component of $\boldsymbol{\theta}$ be denoted by θ^r. The first log-likelihood derivative with respect to θ^r is then defined as $l_r=\partial l(\boldsymbol{\theta})/\partial\theta^r$. If $R_1=r_{11}\ldots r_{1n_1},\ldots\ldots,R_m=r_{m1\ldots mn_m}$ are sets of coordinate indices, joint moments of the log-likelihood derivatives can be defined by $\upsilon_{R_1,\ldots,R_m}=E[l_{R_l}\ldots l_{R_m}]$, where, for example, $\upsilon_{s,tu}=E[l_s l_{tu}]$ and $\upsilon_{a,b,c,de}=E[l_a l_b l_c l_{de}]$.

The expected information, known as the Fisher information, is defined by $i_{rs}=E[l_r l_s]$ where the indices r, s are arbitrary. Lifting the indices produces quantities that are denoted by $\upsilon^{R_1,\ldots,R_m}=i^{r_{11}s_{11}}\ldots i^{r_{mn_m}s_{mn_m}}\upsilon_{s_{11}\ldots s_{1n_1},\ldots,s_{m1}\ldots s_{mn_m}}$, where $i^{rs}=[\mathbf{i}^{-1}]_{rs}$ is the r, s component of the inverse $\mathbf{i}^{-1}$ of the expected information matrix $\mathbf{i}$. The inverse of the Fisher information matrix $\mathbf{i}^{-1}$ is also known as the Cramer–Rao lower bound (CRLB). Here, as elsewhere, the Einstein summation convention is used. That is, if an index occurs twice in a term, once in the subscript and once in the superscript, summation over the index is implied.

The MLE $\hat{\boldsymbol{\theta}}$, the value of $\boldsymbol{\theta}$ that maximizes $l(\boldsymbol{\theta})$ for the given data $\mathbf{X}$,[1,6,7] can now be expressed as an asymptotic expansion around $\boldsymbol{\theta}$ in increasing orders of inverse sample size n^{-1} or equivalently SNR. Following the derivation of Barndorff-Nielsen and Cox,[7] the component l_r is first expanded around $\boldsymbol{\theta}$ as

$$\hat{l}_r=l_r+l_{rs}(\hat{\theta}-\theta)^s+\tfrac{1}{2}l_{rst}(\hat{\theta}-\theta)^s(\hat{\theta}-\theta)^t + \tfrac{1}{6}l_{rstu}(\hat{\theta}-\theta)^s(\hat{\theta}-\theta)^t(\hat{\theta}-\theta)^u+\cdots, \tag{1}$$

where $(\hat{\theta}-\theta)^r=\hat{\theta}^r-\theta^r$. Equation (1) is then inverted to obtain an asymptotic expansion for $(\hat{\theta}-\theta)^r$, as shown in Appendix D. After collecting terms of the same asymptotic order, this can be expressed as[7]

$$(\hat{\theta}-\theta)^r=\underbrace{i^{rs}l_s}_{O_p(n^{-1/2})}+\underbrace{\tfrac{1}{2}\upsilon^{rst}l_s l_t+i^{rs}i^{tu}H_{st}l_u}_{O_A(n^{-1})}+\underbrace{\tfrac{1}{6}(\upsilon^{rstu}+3\upsilon^{rsv}i_{vw}\upsilon^{wtu})l_s l_t l_u}_{O_A(n^{-3/2})} + \underbrace{\upsilon^{rsu}i^{tv}H_{uv}l_s l_t+\tfrac{1}{2}i^{rs}\upsilon^{tuv}H_{st}l_u l_v+\tfrac{1}{2}i^{rs}i^{tu}i^{vw}H_{stv}l_u l_w}_{O_A(n^{-3/2})}+\underbrace{i^{rs}i^{tu}i^{vw}H_{st}H_{uv}l_w}_{O_A(n^{-3/2})}+O_A(n^{-2})+\cdots, \tag{2}$$

where $H_R \equiv l_R - v_R$. The terms are organized in decreasing asymptotic order. The drops occur in asymptotic orders of $n^{-1/2}$ under ordinary repeated sampling, which is equivalent to an asymptotic drop of $(\mathrm{SNR})^{-1/2}$. The asymptotic orders of each set of terms are indicated by symbols such as $O_A(n^{-m})$ which denotes a polynomial that will be of order n^{-m} when n is large but may contain higher order terms, i.e., $O_p(n^{-(m+1)})$, that can be significant when n is small. Here the symbol $O_p(n^{-m})$ denotes a polynomial of exactly order n^{-m} for all values of n.

The first order bias of the MLE is then the expected value of Eq. (2), as derived by Barndorff-Nielsen and Cox,[7]

$$\begin{aligned} b(\hat{\theta}^r) &= E[(\hat{\theta}-\theta)^r] \\ &= \underbrace{\tfrac{1}{2} i^{ra} i^{bc}(v_{abc}+2v_{ab,c})}_{O_p(n^{-1})} + O_p(n^{-3/2}) + \cdots . \end{aligned} \tag{3}$$

Here we take a further step and use Eq. (2) to derive the error correlation of the MLE to second order as given by the order-separated expression

$$\begin{aligned} \mathrm{Cor}(\hat{\theta}^r,\hat{\theta}^a) &= E[(\hat{\theta}-\theta)^r(\hat{\theta}-\theta)^a] \\ &= \underbrace{[i^{ra}]}_{O_p(n^{-1})} + [\underbrace{2 i^{mb} i^{nc} v_{lmn}(i^{rs}i^{la}+i^{as}i^{lr}) v_{s,b,c}(n^1) + \tfrac{1}{2} i^{cd} i^{ef}(i^{rs}i^{ab}+i^{as}i^{rb}) v_{bce,d,f,s}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ i^{tu}(i^{rs}i^{ab}i^{cd}+i^{rd}i^{ab}i^{cs}+i^{ad}i^{rb}i^{cs}) v_{st,u,bc,d}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ i^{bm}i^{cq}i^{tp} v_{lmn} v_{opq}(\tfrac{1}{4} i^{rl}i^{ao}i^{sn} + \tfrac{1}{2} i^{rs}i^{al}i^{on} + \tfrac{1}{2} i^{as}i^{rl}i^{on}) v_{s,t,b,c}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ \tfrac{1}{2} i^{sm} v_{lmn}(i^{tn}i^{cd}(i^{rl}i^{ab}+i^{al}i^{rb}) + 2 i^{bn}i^{cd}(i^{rl}i^{at}+i^{al}i^{rt}) + i^{cl}i^{tn}(i^{rd}i^{ab}+i^{ad}i^{rb})) v_{s,t,bc,d}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ \tfrac{1}{6} i^{mb}i^{nc}i^{od} v_{lmno}(i^{rs}i^{la}+i^{as}i^{rl}) v_{s,b,c,d}(n^2) + 4 i^{bm}(i^{rs}i^{al}+i^{as}i^{rl}) v_{s,m,lb}(n^1)}_{O_p(n^{-2})}] + O_p(n^{-3}) + \cdots , \end{aligned} \tag{4}$$

where notation such as $v_{bce,d,f,s}(n^2)$ means the n^2 order terms of the joint moment $v_{bce,d,f,s}$.

Using the identity $\mathrm{Cov}(\hat{\theta}^r,\hat{\theta}^a) = \mathrm{Cor}(\hat{\theta}^r,\hat{\theta}^a) - b(\theta^r)b(\theta^a)$, we obtain the following expression for the covariance of the MLE to second order:

$$\begin{aligned} \mathrm{Cov}(\hat{\theta}^r,\hat{\theta}^a) &= E[(\hat{\theta}^r - E[\theta^r])(\hat{\theta}^a - E[\theta^a])] \\ &= \underbrace{[i^{ra}]}_{O_p(n^{-1})} + [\underbrace{2 i^{mb} i^{nc} v_{lmn}(i^{rs}i^{la}+i^{as}i^{lr}) v_{s,b,c}(n^1) + \tfrac{1}{2} i^{cd} i^{ef}(i^{rs}i^{ab}+i^{as}i^{rb}) v_{bce,d,f,s}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ i^{tu}(i^{rs}i^{ab}i^{cd}+i^{rd}i^{ab}i^{cs}+i^{ad}i^{rb}i^{cs}) v_{st,u,bc,d}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ i^{bm}i^{cq}i^{tp} v_{lmn} v_{opq}(\tfrac{1}{4} i^{rl}i^{ao}i^{sn} + \tfrac{1}{2} i^{rs}i^{al}i^{on} + \tfrac{1}{2} i^{as}i^{rl}i^{on}) v_{s,t,b,c}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ \tfrac{1}{2} i^{sm} v_{lmn}(i^{tn}i^{cd}(i^{rl}i^{ab}+i^{al}i^{rb}) + 2 i^{bn}i^{cd}(i^{rl}i^{at}+i^{al}i^{rt}) + i^{cl}i^{tn}(i^{rd}i^{ab}+i^{ad}i^{rb})) v_{s,t,bc,d}(n^2)}_{O_p(n^{-2})} \\ &\quad \underbrace{+ \tfrac{1}{6} i^{mb}i^{nc}i^{od} v_{lmno}(i^{rs}i^{la}+i^{as}i^{rl}) v_{s,b,c,d}(n^2) + 4 i^{bm}(i^{rs}i^{al}+i^{as}i^{rl}) v_{s,m,lb}(n^1)}_{O_p(n^{-2})} \\ &\quad \underbrace{- \tfrac{1}{4} i^{rs}i^{tu}i^{aw}i^{yz}(v_{stu}v_{wyz} + 2 v_{stu} v_{wy,z}(n^1) + 2 v_{st,u} v_{wyz}(n^1) + 4 v_{st,u} v_{wy,z}(n^1))}_{O_p(n^{-2})} + O_p(n^{-3}) + \cdots . \end{aligned} \tag{5}$$

The first order covariance term i^{ra} is the r, a component of the inverse of the Fisher information, or the r, a component of the CRLB. A bound on the lowest possible mean square error of an *unbiased scalar* estimate that involves inverse sample size orders higher than n^{-1} was introduced by Bhattacharyya.[2] While it involves derivatives of the likelihood function, it is quite different from the multivariate covariance derived in Eq. (5) that is valid for multivariate estimates that may be biased. For discrete random variables, expressions equivalent to Eqs. (3)–(5) have been obtained in a significantly different form via a different approach by Bowman and Shenton.[14]

A *necessary* condition for the MLE to become asymptotically unbiased. This is for the first order bias of Eq. (1) to become much smaller than the true value of the parameter θ^r. Similarly, a *necessary* condition for the MLE to asymptotically attain minimum variance is for the sum of second order terms in Eq. (5) to become much smaller than the first order term, which is the CRLB.

III. ASYMPTOTIC BIAS, ERROR CORRELATION AND COVARIANCE OF THE MLE FOR GAUSSIAN DATA

The asymptotic expressions presented for the bias, error correlation, and covariance of the MLE in Sec. II are now evaluated for real multivariate Gaussian data. General multivariate Gaussian data can be described by the conditional probability density

$$p(\mathbf{X};\boldsymbol{\theta})=\frac{1}{(2\pi)^{nN/2}}\frac{1}{|\mathbf{C}(\boldsymbol{\theta})|^{n/2}}\exp\Bigg\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{X}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T \times\mathbf{C}^{-1}(\boldsymbol{\theta})(\mathbf{X}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\Bigg\}, \tag{6}$$

where the data $\mathbf{X}=[\mathbf{X}_1^T\mathbf{X}_2^T\mathbf{X}_3^T\ldots\mathbf{X}_n^T]^T$ are comprised of n independent and identically distributed N-dimensional data vectors $\mathbf{X}_i$ to show an explicit dependence under normal repeated sampling for convenient reference. It is noteworthy that the CRLB is always proportional to $1/n$ but may be proportional to a more complicated function of the length of the data vector N.

We begin by deriving the first order bias for the general multivariate Gaussian case where the data covariance $\mathbf{C}$ and the data mean $\boldsymbol{\mu}$ depend on the parameter vector $\boldsymbol{\theta}$. The joint moments required to evaluate both the error correlation and covariance for the general case are quite complicated but not of great relevance in most standard ocean acoustic and signal processing problems.[8] They are not derived in this paper, but are the subject of another work where the second order bias is also derived.[15] We instead define two special cases that have great practical value, since they describe a deterministic signal in additive noise and a fully randomized signal in noise, respectively. In the former the data covariance $\mathbf{C}$ is independent of the parameter vector $\boldsymbol{\theta}$, while the mean $\boldsymbol{\mu}$ depends on $\boldsymbol{\theta}$ which is the subject of the estimation problem. In the latter, the data mean $\boldsymbol{\mu}$ is zero while the covariance $\mathbf{C}$ depends on the parameter vector $\boldsymbol{\theta}$ to be estimated. In the latter case, the sample covariance of the data is a sufficient statistic that contains all measurement information about the parameters to be estimated.[1,16]

The assumption of Gaussian data is valid, by virtue of the central limit theorem even for small n and N, when the total received field is the sum of a large number of statistically independent contributions. In the case of a deterministic signal in additive noise, the additive noise typically arises from a large number of independent sources distributed over the sea surface.[17] These noise sources may be either caused by the natural action of wind and waves on the sea surface, or they may be generated by ocean-going vessels.[18]

A particular fully randomized Gaussian signal model that is very widely used and enjoys a long history in acoustics, optics, and radar[19,20] is the circular complex Gaussian random (CCGR) model. The basic assumption in this model is that at any time instant, the received signal field is a CCGR variable.[9,19] This means that the real and imaginary parts of the instantaneous field are independent and identically distributed zero-mean Gaussian random variables. In active detection and imaging problems, this model is typically used to describe scattering from fluctuating targets[21,22] and surfaces with wavelength scale roughness.[9] When the target or resolved surface patch is large compared to the wavelength, the total received field can be thought of as arising from the sum of a large number of independent scatters so that the central limit theorem applies. Since World War II, the CCGR signal model has been used to describe ocean-acoustic transmission scintillation in what is known as the saturated region of multi-modal propagation.[19,23,24] In this regime, natural disturbances in the waveguide, such as underwater turbulence and passing surface or internal gravity waves, lead to such randomness in the medium that the waveguide modes at the receiver can be treated as statistically independent entities. The central limit theorem can then be invoked for the total received field, which behaves as a CCGR process in time.[16,19] In passive source localization problems in ocean-acoustics, the source signal is typically mechanical noise that is accidentally radiated into the ocean by a vessel. This noise typically has both narrow and broadband components that arise from a broad distribution of independent mechanical interactions that lead to a signal that can be represented as a CCGR process in time. The CCGR signal model has become a very standard model in ocean-acoustic matched field processing.[16,25,26]

A. The general multivariate Gaussian case

We obtain the following expression for the first order bias of the MLE given general multivariate Gaussian data

$$b(\hat{\theta}^r)=\tfrac{1}{2}i^{rs}i^{tu}(\upsilon_{stu}+2\upsilon_{st,u})+O_p(n^{-3/2})$$

$$\begin{aligned}=\sum_{s=1}^{m}\sum_{t=1}^{m}\sum_{u=1}^{m}\frac{2}{n}[\mathbf{i}^{-1}]_{rs}[\mathbf{i}^{-1}]_{tu}\Bigg\{&\frac{1}{2}\mathrm{tr}\left(\mathbf{C}^{-1}\frac{\partial^2\mathbf{C}}{\partial\theta^s\,\partial\theta^t}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial\theta^u}\right)+\left(\frac{\partial^2\boldsymbol{\mu}}{\partial\theta^s\,\partial\theta^t}\right)^T\mathbf{C}^{-1}\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta^u}\right)+\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta^s}\right)^T\mathbf{C}^{-1}\left(\frac{\partial\mathbf{C}}{\partial\theta^u}\right)\mathbf{C}^{-1}\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta^t}\right)\\&-\sum_{s,t}\left(\frac{\partial^2\boldsymbol{\mu}}{\partial\theta^s\,\partial\theta^u}\right)^T\mathbf{C}^{-1}\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta^t}\right)-\sum_{s,t}\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta^s}\right)^T\mathbf{C}^{-1}\left(\frac{\partial\mathbf{C}}{\partial\theta^t}\right)\mathbf{C}^{-1}\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta^u}\right)-\frac{1}{2}\sum_{s,t}\mathrm{tr}\left(\mathbf{C}^{-1}\frac{\partial^2\mathbf{C}}{\partial\theta^s\,\partial\theta^u}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial\theta^t}\right)\Bigg\},\end{aligned} \tag{7}$$

by substituting Eqs. (A1)–(A3) for the relevant joint moments into Eq. (3) for the first order bias, where $\Sigma_{s,t}$ indicates a sum over all possible permutations of *s*, *t* orderings, a total of two. For example, $\Sigma_{s,t} v_{st} = v_{st} + v_{ts}$.

It should be noted that the expression contains both tensor notation, denoted by the indices *s*, *t*, and *u*, and vector-matrix notation. For the first order bias, only first and second order parameter derivatives are required of the mean and covariance.

Suppose, for example, the bias of the vector

$$\hat{\theta} = \begin{bmatrix} \hat{\mu} \\ \hat{C} \end{bmatrix}$$

is desired, where $\hat{\mu}$ and $\hat{C}$ are the maximum likelihood estimates of the mean and variance, respectively, from a set of n independent and identically distributed Gaussian random variables x_i. The bias obtained from Eq. (6) is zero for the mean component and $-(C/n)$ for the variance component. This result can be readily verified by taking expectation values directly.[27]

It is noteworthy that the first order bias of a *scalar* parameter estimate always vanishes for general Gaussian data as can be seen by inspection of Eq. (7).

B. Deterministic signal in additive noise, parameter-independent covariance

The multivariate error correlation and covariance of the MLE can be obtained to second order for a deterministic signal vector in additive Gaussian noise by substituting Eqs. (B1)–(B10) into Eqs. (4) and (5), respectively. In this case, **C** is independent of $\boldsymbol{\theta}$ in Eq. (6). For scalar parameter and data, the following simple expressions for the mean-square error and variance are obtained

$$\mathrm{MSE}(\hat{\theta}) = \underbrace{\frac{C}{n(\mu')^2}}_{O_p(n^{-1})} + \underbrace{\frac{15C^2(\mu'')^2}{4n^2(\mu')^6} - \frac{C^2(\mu''')}{n^2(\mu')^5}}_{O_p(n^{-2})}, \tag{8}$$

$$\mathrm{var}(\hat{\theta}) = \underbrace{\frac{C}{n(\mu')^2}}_{O_p(n^{-1})} + \underbrace{\frac{14C^2(\mu'')^2}{4n^2(\mu')^6} - \frac{C^2(\mu''')}{n^2(\mu')^5}}_{O_p(n^{-2})}. \tag{9}$$

Suppose, for example, that the bias, mean-square error and variance of the MLE of the parameter $\theta = \mu^2$ are desired, where the x_i are again n independent and identically distributed Gaussian random variables, and C is independent of θ. The corresponding bias C/n, mean-square error $4C\mu^2/n + 3C^2/n^2$, and variance $4C\mu^2/n + 2C^2/n^2$, obtained using Eqs. (3)–(5) can be readily shown to correspond to those obtained by taking expectation values directly.[27] Since the MLE for $\theta = \mu^2$ is biased, the Bhattacharyya bound does not hold for this example and in fact can exceed the actual variance of the MLE.[27]

C. Random signal in noise: Zero-mean and parameter-dependent covariance

Similarly, the error correlation and covariance of the MLE can be obtained to second order for a zero-mean Gaussian random signal vector in Gaussian noise by substituting Eqs. (C1)–(C10) into Eqs. (4) and (5), respectively. In this case, $\boldsymbol{\mu}$ is zero in Eq. (6). For scalar parameter and data, the following simple expressions for the mean-square error and variance expressions are obtained:

$$\mathrm{MSE}(\hat{\theta}) = \underbrace{\frac{2C^2}{n(C')^2}}_{O_p(n^{-1})} - \underbrace{\frac{8C^3(C'')}{n^2(C')^4} + \frac{15C^4(C'')^2}{n^2(C')^6} - \frac{4C^4(C''')}{n^2(C')^5}}_{O_p(n^{-2})}, \tag{10}$$

$$\mathrm{var}(\hat{\theta}) = \underbrace{\frac{2C^2}{n(C')^2}}_{O_p(n^{-1})} - \underbrace{\frac{8C^3(C'')}{n^2(C')^4} + \frac{14C^4(C'')^2}{n^2(C')^6} - \frac{4C^4(C''')}{n^2(C')^5}}_{O_p(n^{-2})}. \tag{11}$$

Suppose, for example, that the bias, mean-square error, and variance of the MLE of the parameter $\theta = C^2$ are desired, where the x_i are n independent and identically distributed Gaussian random variables with zero-mean. It can be readily shown that the corresponding bias $2C^2/n$, mean-square error $8C^4/n + 44C^4/n^2$, and variance $8C^4/n + 40C^4/n^2$, obtained using Eqs. (3)–(5), correspond to those obtained by taking expectation values directly.[27]

IV. CONTINUOUS GAUSSIAN DATA: SIGNAL EMBEDDED IN WHITE GAUSSIAN NOISE

Let a real signal $\mu(t;\theta)$ that depends on parameter θ be received together with uncorrelated white Gaussian noise of power spectral density $N_0/2$ that is independent of θ. Suppose the real signal has Fourier transform $\mu(t;\theta) \leftrightarrow \Psi(f;\theta)$. The complex analytic signal and its Fourier transform $\tilde{\mu}(t;\theta) \leftrightarrow \tilde{\Psi}(f;\theta)$ are conventionally defined such that $\tilde{\Psi}(f;\theta) = 2\Psi(f;\theta)$ for $f>0$, $\tilde{\Psi}(f;\theta) = 0$ for $f<0$, and $\tilde{\Psi}(f;\theta) = \Psi(f;\theta)$ for $f=0$, so that $\mu(t;\theta) = \mathrm{Re}\{\tilde{\mu}(t;\theta)\}$. The total received analytic signal, $\tilde{\varphi}_i(t)$, then follows the conditional probability density[28]

$$p(\tilde{\varphi}_i(t);\theta) = k \exp\left\{ -\frac{1}{2N_0} \int_0^T |(\tilde{\varphi}_i(t) - \tilde{\mu}(t;\theta))|^2 \, dt \right\}, \tag{12}$$

where k is a normalization constant. The bias, the mean-square error, and the variance of the MLE $\hat{\theta}$ are obtained from Eqs. (3)–(5) as

$$b(\hat{\theta}) = -\frac{N_0}{2} \frac{\mathrm{Re}\{\tilde{I}_2\}}{\tilde{I}_1^2}, \tag{13}$$

$$\mathrm{MSE}(\hat{\theta})=\underbrace{\frac{N_0}{\tilde{I}_1}}_{O_p(n^{-1})}+\underbrace{\frac{15N_0^2\,\mathrm{Re}\{\tilde{I}_2\}^2}{4\tilde{I}_1^4}-\frac{N_0^2\,\mathrm{Re}\{\tilde{I}_3\}}{\tilde{I}_1^3}}_{O_p(n^{-2})}, \tag{14}$$

$$\mathrm{var}(\hat{\theta})=\underbrace{\frac{N_0}{\tilde{I}_1}}_{O_p(n^{-1})}+\underbrace{\frac{14N_0^2\,\mathrm{Re}\{\tilde{I}_2\}^2}{4\tilde{I}_1^4}-\frac{N_0^2\,\mathrm{Re}\{\tilde{I}_3\}}{\tilde{I}_1^3}}_{O_p(n^{-2})}, \tag{15}$$

after evaluating the joint moments for the parameter-independent covariance, where $\tilde{I}_1$, $\tilde{I}_2$, $\tilde{I}_3$ are defined as follows:

$$\tilde{I}_1=\int\left(\frac{\partial\tilde{\mu}(t;\theta)}{\partial\theta}\right)^*\left(\frac{\partial\tilde{\mu}(t;\theta)}{\partial\theta}\right)dt, \tag{16}$$

$$\tilde{I}_2=\int\left(\frac{\partial\tilde{\mu}(t;\theta)}{\partial\theta}\right)^*\left(\frac{\partial\tilde{\mu}^2(t;\theta)}{\partial\theta^2}\right)dt, \tag{17}$$

$$\tilde{I}_3=\int\left(\frac{\partial\tilde{\mu}(t;\theta)}{\partial\theta}\right)^*\left(\frac{\partial\tilde{\mu}^3(t;\theta)}{\partial\theta^3}\right)dt. \tag{18}$$

There are two important issues to note. First, we are now working with continuously measured data as opposed to the discrete data vectors of Sec. III. Second, the fact that we are only estimating a scalar rather than a vector parameter greatly simplifies the evaluation of the joint moments.

V. TIME-DELAY ESTIMATION

Suppose $\tilde{\mu}(t;\theta)=\tilde{\mu}(t-\tau)$ in Eq. (12) so that the scalar time delay $\theta=\tau$ is to be estimated. The MLE $\hat{\theta}=\hat{\tau}$ of time-delay τ corresponds to the peak output of a matched filter for a signal received in additive Gaussian noise.[8] Estimates of the time delay between transmitted and received signal waveforms are typically used in active-sonar and radar applications to determine the range of a target in a nondispersive medium. The asymptotic bias, mean-square error, and variance of $\hat{\tau}$ are obtained by substituting τ for θ in Eqs. (13)–(18).

The following alternative expressions are obtained for Eqs. (16)–(18) by applying Parseval's Theorem

$$\tilde{I}_1=\int\left(\frac{\partial\tilde{\mu}(t-\tau)}{\partial\tau}\right)^*\left(\frac{\partial\tilde{\mu}(t-\tau)}{\partial\tau}\right)dt$$
$$=(2\pi)^2\int_0^\infty f^2|\tilde{\Psi}(f)|^2\,df, \tag{19}$$

$$\tilde{I}_2=\int\left(\frac{\partial\tilde{\mu}(t-\tau)}{\partial\tau}\right)^*\left(\frac{\partial\tilde{\mu}^2(t-\tau)}{\partial\tau^2}\right)dt$$
$$=j(2\pi)^3\int_0^\infty f^3|\tilde{\Psi}(f)|^2\,df, \tag{20}$$

$$\tilde{I}_3=\int\left(\frac{\partial\tilde{\mu}(t-\tau)}{\partial\tau}\right)^*\left(\frac{\partial\tilde{\mu}^3(t-\tau)}{\partial\tau^3}\right)dt$$
$$=-(2\pi)^4\int_0^\infty f^4|\tilde{\Psi}(f)|^2\,df. \tag{21}$$

Noting that the first order bias of Eq. (13) is directly proportional to $\mathrm{Re}\{\tilde{I}_2\}$, where $\mathrm{Re}\{\tilde{I}_2\}=0$ from Eq. (19), we find that *the first order bias for the maximum likelihood time-delay estimation problem is identically zero*, as expected by inspection of Eq. (7).

To evaluate the mean-square error and the variance, all three integrals are needed. Noting that $\mathrm{Re}\{\tilde{I}_2\}=0$ and $\mathrm{Re}\{\tilde{I}_3\}=\tilde{I}_3$ the following results are obtained:

$$\mathrm{MSE}(\hat{\tau})=\mathrm{var}(\hat{\tau})=\underbrace{\frac{N_0}{\tilde{I}_1}}_{O_p(n^{-1})}-\underbrace{\left(\frac{N_0}{\tilde{I}_1}\right)^2\frac{\mathrm{Re}\{\tilde{I}_3\}}{\tilde{I}_1}}_{O_p(n^{-2})}+O_p(n^{-3}), \tag{22}$$

where $\tilde{I}_1$, $\tilde{I}_3$ are evaluated in Eqs. (19) and (21), respectively. The mean-square error can also be expressed explicitly in terms of $\mathrm{SNR}=2E/N_0$ and signal parameters via

$$\mathrm{MSE}(\hat{\tau})=\mathrm{var}(\hat{\tau})$$
$$=\underbrace{\frac{1}{(2E/N_0)(\beta^2+\omega_c^2)}}_{O_p(n^{-1})}$$
$$+\underbrace{\frac{1}{(2E/N_0)^2}\frac{\gamma^4+8\omega_c^2\beta^2+3\omega_c^4}{(\beta^2+\omega_c^2)^3}}_{O_p(n^{-2})}+O_p(n^{-3}), \tag{23}$$

where $f_c=\omega_c/(2\pi)$ is the carrier frequency, E is the total energy of the real signal, β is commonly defined as the signal's root mean square (rms) bandwidth, and γ^4 is the fourth moment of the expected signal's energy spectrum.

$$2E=\int_{-f_c}^\infty|\tilde{\Psi}(v+f_c)|^2\,dv, \tag{24}$$

$$\beta^2=\frac{(2\pi)^2\int_{-f_c}^\infty v^2|\tilde{\Psi}(v+f_c)|^2\,dv}{2E}, \tag{25}$$

$$\gamma^4=\frac{(2\pi)^4\int_{-f_c}^\infty v^4|\tilde{\Psi}(v+f_c)|^2\,dv}{2E}. \tag{26}$$

Equation (23) explicitly shows the asymptotic dependence of the MLE time-delay variance on increasing orders of $(\mathrm{SNR})^{-1}$. For a base-banded signal, where $\omega_c=0$, the first order variance term of Eq. (23) is proportional to the inverse of β^2, while the second term is proportional to the ratio γ^4/β^6. While it is well known that the first order variance or CRLB decreases with increasing rms bandwidth at fixed SNR, the behavior of the second order variance term has a more complicated interpretation since it involves both γ and β.

The time-delay MLE asymptotically attains the CRLB when $(\gamma^4+8\omega_c^2\beta^2+3\omega_c^4)/(\beta^2+\omega_c^2)^2 \ll 2E/N_0 = \mathrm{SNR}$. For a base-banded signal, this condition means that the SNR must be much larger than the *kurtosis* γ^4/β^4 of the expected signal's energy spectrum. This can be interpreted as meaning that as the signal's energy spectrum becomes more peaked, higher SNR is necessary to attain the CRLB.

Example 1

Assume a real Gaussian base-banded signal with a constant energy. Its waveform can be represented as

$$h(t)=\frac{1}{\sqrt{\tau_s}}\exp(-\pi(t^2/\tau_s^2)), \quad |t|\leqslant T/2,$$

where $h\leftrightarrow H$. For real signals, $2E$ is replaced by $E/2$, and Ψ replaces $\tilde{\Psi}$ in Eqs. (24)–(26), where in the present case $\Psi = H$. Under the assumption that τ_s/T is sufficiently small that the limits of integration $[-T/2,T/2]$ can be well approximated as $[-\infty,\infty]$, Eq. (23) for the variance of the time-delay MLE can be written to second order as

$$\mathrm{var}(\hat{\tau})=\underbrace{\frac{N_0}{2}\frac{\sqrt{2}}{\pi}\tau_s^2}_{O_p(n^{-1})}+\underbrace{\left(\frac{N_0}{2}\right)^2\frac{6}{\pi}\tau_s^2}_{O_p(n^{-2})}.$$

Since the signal's energy is $1/\sqrt{2}$, the first and second order terms equalize when the SNR is 3, which is the kurtosis of a Gaussian density, where the $\mathrm{SNR}=2E/N_0$. This makes sense because a Gaussian signal has a Gaussian energy spectrum by the convolution theorem. For SNRs less than 3, or in decibels for $10\log \mathrm{SNR}<5$ dB, the second order term is higher than the first and the CRLB is a poor estimate of the true mean-square error. Moreover, since $1/\tau_s$ is a measure of the signal's bandwidth, decreasing τ_s, or increasing the signal's bandwidth, will decrease both first and second order variance terms, and so improve the time-delay estimate.

VI. DOPPLER SHIFT ESTIMATION

Suppose now that a *narrow band* signal waveform is transmitted in a nondispersive medium and measured with additive Gaussian noise at a receiver that is moving relative to the source at low Mach number $u/c\ll 1$, where u is the speed of relative motion and c the speed of wave propagation. The expected analytic signal waveform at the receiver $\tilde{\mu}(t;f_D)$ is then frequency shifted with Doppler-shift parameter $f_D=-2u/c$. The total signal and noise measured at the receiver will then obey the conditional probability density of Eq. (12) with $\theta=f_D$. The goal now is to examine the asymptotic statistics of the MLE $\hat{f}_D$ for the Doppler-shift parameter.

With the given assumptions, the signal waveform can be represented as

$$\tilde{\mu}(t;f_D)=\tilde{g}(t)e^{j2\pi t(f_c+f_D)}, \tag{27}$$

where the complex envelope $\tilde{g}(t)$ is known and is zero outside the interval $-T/2\leqslant t\leqslant T/2$. By applying Parseval's Theorem, the following expressions are obtained:

$$\tilde{I}_1=\int\left(\frac{\partial\tilde{\mu}(t;f_D)}{\partial f_D}\right)^*\left(\frac{\partial\tilde{\mu}(t;f_D)}{\partial f_D}\right)dt$$

$$=(2\pi)^2\int_{-T/2}^{T/2}t^2|\tilde{g}(t)|^2\,dt, \tag{28}$$

$$\tilde{I}_2=\int\left(\frac{\partial\tilde{\mu}(t;f_D)}{\partial f_D}\right)^*\left(\frac{\partial\tilde{\mu}^2(t;f_D)}{\partial f_D^2}\right)dt$$

$$=j(2\pi)^3\int_{-T/2}^{T/2}t^3|\tilde{g}(t)|^2\,dt, \tag{29}$$

$$\tilde{I}_3=\int\left(\frac{\partial\tilde{\mu}(t;f_D)}{\partial f_D}\right)^*\left(\frac{\partial\tilde{\mu}^3(t;f_D)}{\partial f_D^3}\right)dt$$

$$=-(2\pi)^4\int_{-T/2}^{T/2}t^4|\tilde{g}(t)|^2\,dt. \tag{30}$$

To evaluate the bias, $\tilde{I}_1$ and $\mathrm{Re}\{\tilde{I}_2\}$ are substituted into Eq. (13). Noting that $\mathrm{Re}\{\tilde{I}_2\}=0$, we find that *the first order bias for the Doppler-shift MLE* $\hat{f}_D$ *is identically zero*, as expected by inspection of Eq. (7).

Equations (28)–(30) are then substituted into Eqs. (14)–(15) to evaluate the mean-square error and the variance to second order. The resulting relations for the second order mean-square error and variance of the Doppler-shift MLE are similar to those obtained for the time-delay MLE:

$$\mathrm{MSE}(\hat{f}_D)=\mathrm{var}(\hat{f}_D)$$

$$=\underbrace{\frac{N_0}{\tilde{I}_1}}_{O_p(n^{-1})}-\underbrace{\left(\frac{N_0}{\tilde{I}_1}\right)^2\frac{\mathrm{Re}\{\tilde{I}_3\}}{\tilde{I}_1}}_{O_p(n^{-2})}+O_p(n^{-3}), \tag{31}$$

since the two problems are related through the time-frequency duality principle. Expressing $\tilde{I}_1$, $\tilde{I}_3$ in terms of SNR and signal parameters then explicitly yields the Doppler-shift MLE mean-square error in terms of increasing orders of $(\mathrm{SNR})^{-1}$ as

$$\mathrm{MSE}(\hat{f}_D)=\mathrm{var}(\hat{f}_D)$$

$$=\underbrace{\frac{1}{(2E/N_0)\alpha^2}}_{O_p(n^{-1})}+\underbrace{\frac{1}{(2E/N_0)^2}\frac{\delta^4}{\alpha^6}}_{O_p(n^{-2})}+O_p(n^{-3}), \tag{32}$$

where

$$2E=\int_{-T/2}^{T/2}|\tilde{g}(t)|^2\,dt, \tag{33}$$

$$\alpha^2=\frac{(2\pi)^2\int_{-T/2}^{T/2}t^2|\tilde{g}(t)|^2\,dt}{2E}, \tag{34}$$

$$\delta^4=\frac{(2\pi)^4\int_{-T/2}^{T/2}t^4|\tilde{g}(t)|^2\,dt}{2E}. \tag{35}$$

The Doppler-shift MLE then asymptotically attains the CRLB when $\delta^4/\alpha^4\ll 2E/N_0=\mathrm{SNR}$. For an analytic signal

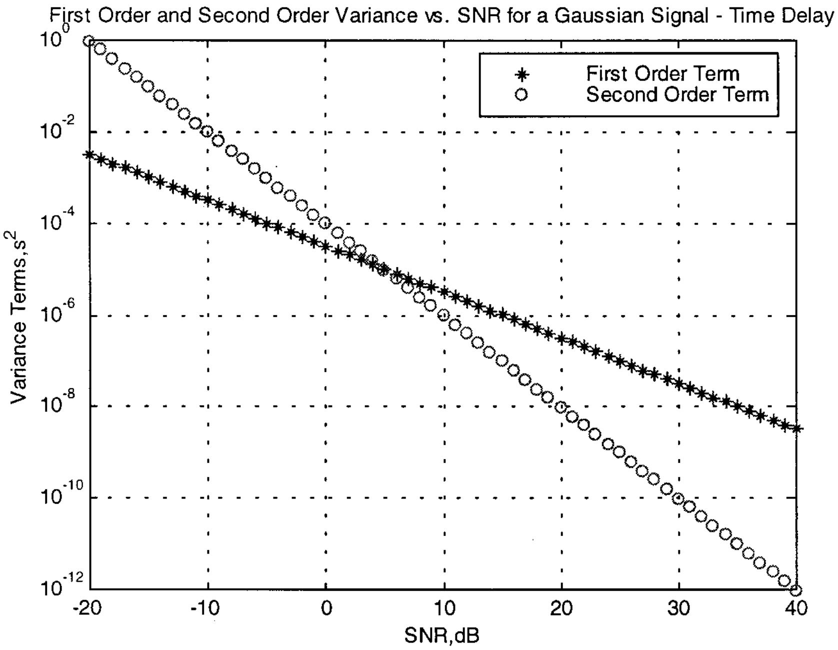

FIG. 1. Gaussian signal time-delay variance terms as a function of SNR.

with symmetric magnitude, this can be interpreted as meaning that the SNR must be large compared to the *kurtosis* of the signal's squared magnitude.

Example 2

For real signals where $\mu(t;f_D)=g(t)\cos 2\pi t(f_c+f_D)$, Eq. (32) can be used when $2E$ is replaced by $4E$ in Eqs. (34)–(35), and the real signal envelope $g(t)$ replaces $\tilde{g}(t)$ in Eqs. (33)–(35). For the CRLB to be attained in this real signal case, the SNR must be large compared to *twice* the *kurtosis* of the squared magnitude of the real signal envelope, assuming a symmetric magnitude. Computing E, α^2, and δ^4 for the real signal envelope $g(t)=h(t)$ of example 1, by Eq. (32), the variance of the Doppler-shift MLE can be written to second order as

$$\mathrm{var}(\hat{f}_D)=\underbrace{\frac{N_0}{2}\frac{2\sqrt{2}}{\pi}\frac{1}{\tau_s^2}}_{O_p(n^{-1})}+\underbrace{\left(\frac{N_0}{2}\right)^2\frac{24}{\pi}\frac{1}{\tau_s^2}}_{O_p(n^{-2})}+O_p(n^{-3}).$$

The first and second order terms equalize when the SNR is 6, twice the Gaussian kurtosis as expected, where SNR $=2E/N_0$. For SNRs less than 6, or in decibels when $10\log \mathrm{SNR}<7.8$ dB, the second order term is higher than the first, and the CRLB provides a poor estimate of the true MSH. Increasing τ_s decreases the signal's bandwidth and so also decreases both first and second order variance terms, which improves the Doppler-shift estimate.

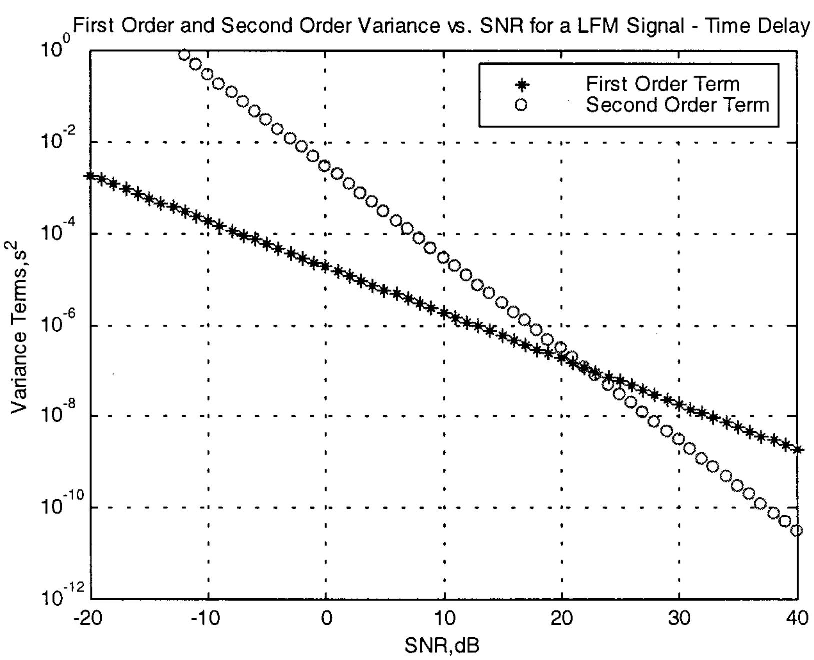

FIG. 2. LFM signal time-delay variance terms as a function of SNR.

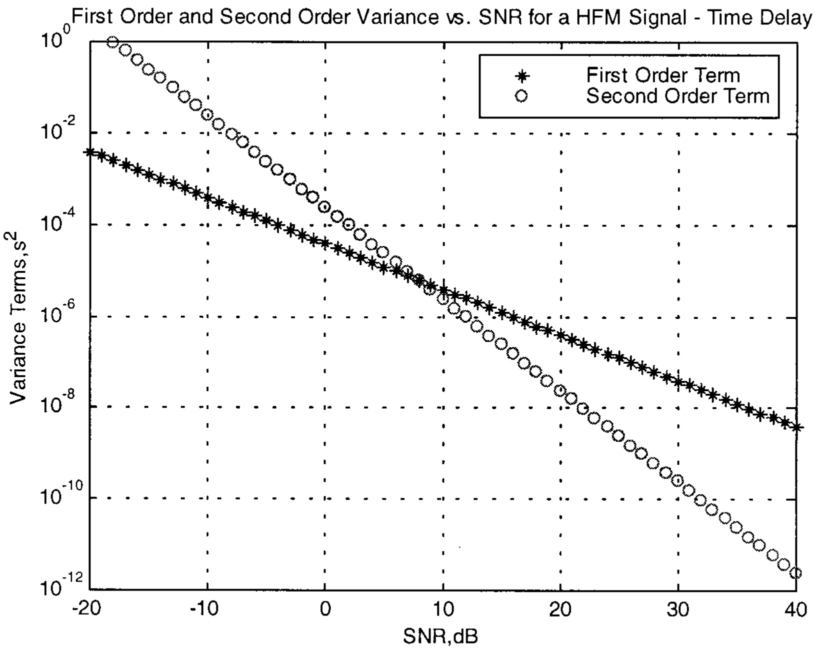

FIG. 3. HFM signal time-delay variance terms as a function of SNR.

VII. ASYMPTOTIC OPTIMALITY OF GAUSSIAN, LFM AND HFM WAVEFORMS IN MAXIMUM LIKELIHOOD TIME-DELAY AND DOPPLER-SHIFT ESTIMATION

The general expressions for the second order variance for both the MLE time-delay and the Doppler-shift estimators, Eqs. (22) and (31), are now implemented for the Gaussian linear frequency modulated (LFM) and hyperbolic frequency modulated (HFM) waveforms. All waveforms are demodulated.

The Gaussian signal is described in examples 1 and 2 above. The LFM signal is defined by

$$\mu(t)=\cos(\omega_0 t+\tfrac{1}{2}bt^2), \quad |t|\leqslant T/2, \tag{36}$$

where $f_0=\omega_0/2\pi$ is the carrier frequency and the bandwidth is given by $bT/2\pi$. The signal is demodulated when multiplied by $\cos(\omega_0 t)$ and low-pass filtered. The HFM signal is defined by

$$\mu(t)=\sin(\alpha \log(1-k(t+T/2))), \quad |t|\leqslant T/2, \tag{37}$$

where $k=(f_2-f_1)/f_2 T$, $\alpha=-2\pi f_1/k$, and f_1 and f_2 are the frequencies that bound the signal's spectrum. This signal is demodulated when multiplied by $\cos(\omega_0 t)$, where $f_0=\sqrt{f_1 f_2}$, and is low-pass filtered. To control sidelobes in the frequency domain, the signal is often multiplied by a temporal window function, or taper. We use the modified Tukey window that has the form

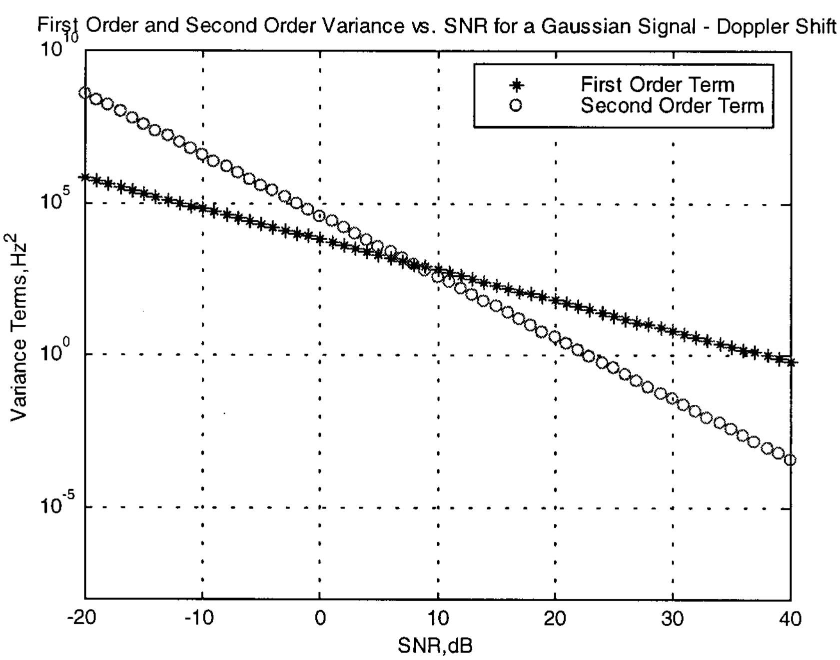

FIG. 4. LFM signal Doppler-shift variance terms as a function of SNR.

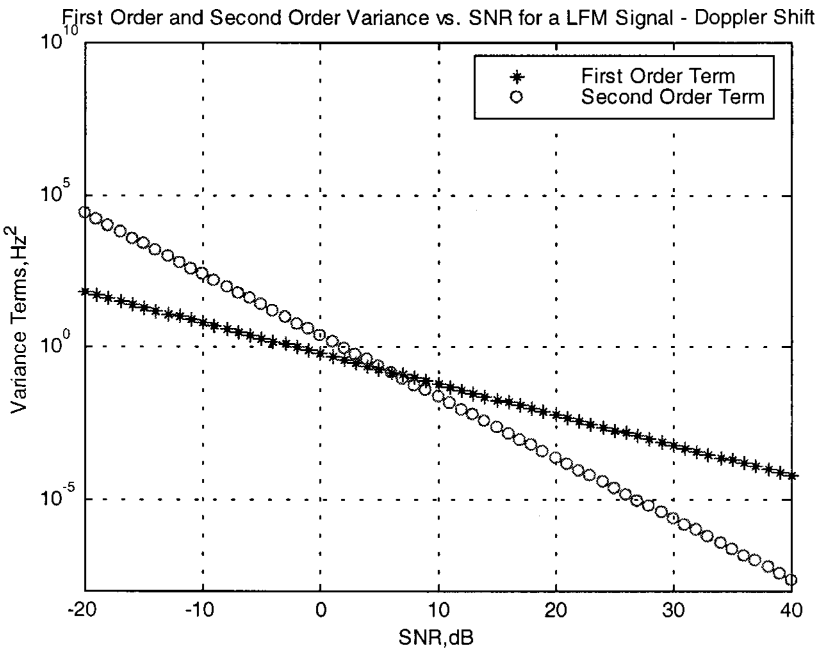

FIG. 5. Gaussian signal Doppler-shift variance terms as a function of SNR.

$$w(t) = \begin{cases} p+(1-p)\sin^2\left(\pi \dfrac{t+T/2}{2T_w}\right) & \text{for } 0\leqslant t\leqslant T_w \\ 1 & \text{for } T_w\leqslant t\leqslant T-T_w \\ p+(1-p)\sin^2\left(\pi \dfrac{(t+T/2)-(T-2T_w)}{2T_w}\right) & \\ & \text{for } T-T_w\leqslant t\leqslant T \end{cases}, \tag{38}$$

where $T_w=0.125T$ is the window duration and $p=0.1$ is the pedestal used.

The dependence of the first and second order variance terms on SNR is presented in Figs. 1–3 for the time-delay MLE and in Figs. 4–6 for the Doppler-shift MLE for the three signals. The bandwidth is fixed at 100 Hz in Figs. 1–6, for a typical low-frequency active-sonar scenario.[29] The figures illustrate some important characteristics of the variance terms for both the time-delay and the Doppler-shift MLE. As SNR increases, the first order variance exhibits the expected linear fall-off and the second order variance falls off with the expected second order power law as can be seen more generally in Eqs. (23) and (32) where the second order term is proportional to $-20\log_{10}(N_0/2E)$, and the first to $-10\log_{10}(N_0/2E)$. The value of either term at a specific bandwidth and SNR can then be used to determine its value at the same bandwidth for all SNRs.

Table I specifies the SNR's values beyond which the second order variance can be neglected relative to the first by

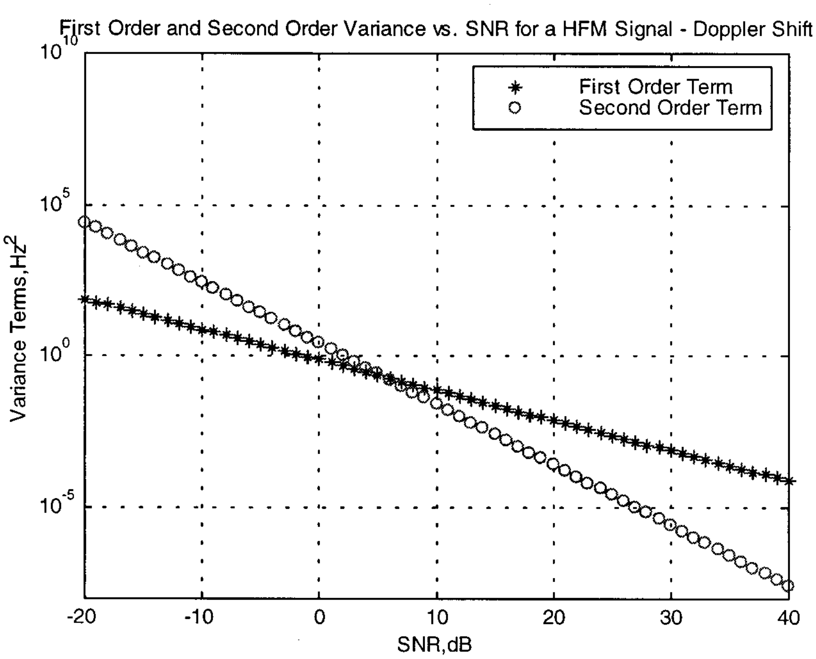

FIG. 6. HFM signal Doppler-shift variance terms as a function of SNR.

TABLE I. Minimum signal-to-noise ratios (SNRs) *necessary* for the MLE to asymptotically attain the CRLB. The given values are the minimum SNRs needed for the CRLB to exceed the second order MLE variance by 10 dB. All signals have 100 Hz bandwidth.

	Gaussian signal	LFM signal	HFM signal
Time Delay	15 dB	32 dB	18 dB
Doppler Shift	18 dB	16 dB	15 dB

showing where the former is an order of magnitude less than the latter. Table I then provides conditions *necessary* for the MLE to attain the CRLB in time-delay and Doppler-shift estimation for the given signals. It also specifies conditions *necessary* for the MLE to be approximated as a linear function of the measured data.

VIII. CONCLUSION

By employing an asymptotic expansion of the likelihood function, expressions for the first order bias, as well as the second order covariance and error correlation of a general MLE, are derived. These expressions are used to determine conditions *necessary* for the MLE to become asymptotically unbiased and attain the CRLB. The approach is then applied to parameter estimation with multivariate Gaussian data. Analytic expressions for the general first order bias of the multivariate Gaussian MLE and the second order error covariance and correlation of the MLE for two special cases of multivariate Gaussian data that are of great practical significance in acoustics, optics, radar, seismology, and signal processing. The first is where the data covariance matrix is independent of the parameters to be estimated, the standard deterministic signal in additive noise scenario. The second is where the data mean is zero and the signal as well as the noise undergo circular complex Gaussian random fluctuations. In a companion paper, the expressions derived here are applied to determine the asymptotic bias, covariance, and mean-square error of maximum likelihood range and depth estimates of a sound source submerged in an ocean waveguide from measured hydrophone array data.[11] Necessary conditions for these source localization estimates to attain the CRLB are also obtained.[11]

In the present paper, general expressions for the first order bias, second order mean-square error, and variance of scalar maximum likelihood time-delay and Doppler-shift estimates are obtained for deterministic signals in additive Gaussian noise. The time-delay MLE is the peak value of a matched filter output. Both time-delay and Doppler-shift MLEs are shown to be unbiased to first order. Analytic conditions on SNR necessary for the time-delay and Doppler-shift MLEs to attain the CRLB are provided in terms of moments of the expected signal's squared magnitude and energy spectrum. For base-banded signals, the time-delay MLE, namely the matched filter estimate, attains the CRLB when the *kurtosis* of the expected signal's energy spectrum is much smaller than the SNR. This can be interpreted as meaning that higher SNR is necessary to attain the CRLB as a demodulated signal's energy spectrum becomes more peaked. The Doppler-shift MLE is found to have dual behavior for narrow band analytic signals.

APPENDIX A: JOINT MOMENTS FOR ASYMPTOTIC GAUSSIAN INFERENCE: GENERAL MULTIVARIATE GAUSSIAN DATA

For the *general multivariate Gaussian* case of Eq. (6), both the mean $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{C}$ depend on the parameter vector $\boldsymbol{\theta}$. The joint moments required to evaluate the first order bias are

$$i_{ab}=\frac{n}{2}\operatorname{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b)+n\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_b\,, \tag{A1}$$

$$\begin{aligned}\upsilon_{abc}=&\frac{n}{3}\sum_{a,b,c}\operatorname{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c)\\&-\frac{n}{4}\sum_{a,b,c}\operatorname{tr}(\mathbf{C}^{-1}\mathbf{C}_{ab}\mathbf{C}^{-1}\mathbf{C}_c)-\frac{n}{2}\sum_{a,b,c}\boldsymbol{\mu}_{ab}^T\mathbf{C}^{-1}\boldsymbol{\mu}_c\\&+\frac{n}{2}\sum_{a,b,c}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\boldsymbol{\mu}_c\,,\end{aligned} \tag{A2}$$

$$\begin{aligned}\upsilon_{ab,c}=&-\frac{n}{2}\sum_{a,b}\operatorname{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c)+n\boldsymbol{\mu}_{ab}^T\mathbf{C}^{-1}\boldsymbol{\mu}_c\\&+\frac{n}{2}\operatorname{tr}(\mathbf{C}^{-1}\mathbf{C}_{ab}\mathbf{C}^{-1}\mathbf{C}_c)\\&-n\sum_{a,b}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\boldsymbol{\mu}_c\,,\end{aligned} \tag{A3}$$

where, for example, $\Sigma_{a,b,c}$ indicates a sum over all possible permutations of a, b and c orderings, leading to a total of six terms. Terms such as $\mathbf{C}_{ab}$ and $\boldsymbol{\mu}_{ab}$ represent the derivatives of the covariance matrix $\mathbf{C}$ and the mean vector $\boldsymbol{\mu}$ with respect to θ^a and θ^b, respectively.

APPENDIX B: JOINT MOMENTS FOR ASYMPTOTIC GAUSSIAN INFERENCE: MULTIVARIATE GAUSSIAN DATA WITH PARAMETER-INDEPENDENT COVARIANCE: DETERMINISTIC SIGNAL IN INDEPENDENT ADDITIVE NOISE

For this case the covariance matrix of Eq. (6) is independent of the parameters to be estimated, i.e., $\partial\mathbf{C}/\partial\theta^i=0$ for all i. The joint moments required to evaluate the first order bias, as well as the second order error correlation and covariance are

$$i_{ab}=n\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_b\,, \tag{B1}$$

$$\upsilon_{abc}(n^1)=-\frac{n}{2}\sum_{a,b,c}\boldsymbol{\mu}_{ab}^T\mathbf{C}^{-1}\boldsymbol{\mu}_c\,, \tag{B2}$$

$$\upsilon_{a,b,c}(n^1)=0, \tag{B3}$$

$$\upsilon_{ab,c}(n^1)=n\boldsymbol{\mu}_{ab}^T\mathbf{C}^{-1}\boldsymbol{\mu}_c\,, \tag{B4}$$

$$\begin{aligned}\upsilon_{abcd}(n^1)=&-\frac{n}{8}\sum_{a,b,c,d}\boldsymbol{\mu}_{ab}^T\mathbf{C}^{-1}\boldsymbol{\mu}_{cd}\\&-\frac{n}{6}\sum_{a,b,c,d}\boldsymbol{\mu}_{abc}^T\mathbf{C}^{-1}\boldsymbol{\mu}_d\,,\end{aligned} \tag{B5}$$

$$v_{a,b,c,d}(n^2)=\frac{n^2}{8}\sum_{a,b,c,d}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_b\boldsymbol{\mu}_c^T\mathbf{C}^{-1}\boldsymbol{\mu}_d, \tag{B6}$$

$$v_{a,b,cd}(n^1)=0, \tag{B7}$$

$$v_{a,b,c,de}(n^2)=\frac{n^2}{2}\sum_{a,b,c}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_b\boldsymbol{\mu}_c^T\mathbf{C}^{-1}\boldsymbol{\mu}_{de}, \tag{B8}$$

$$v_{a,b,cd,ef}(n^2)=\frac{n^2}{2}\sum_{(a,b)\times(cd,ef)}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_{cd}\boldsymbol{\mu}_b^T\mathbf{C}^{-1}\boldsymbol{\mu}_{ef} + n^2\boldsymbol{\mu}_{cd}^T\mathbf{C}^{-1}\boldsymbol{\mu}_{ef}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_b, \tag{B9}$$

$$v_{a,b,c,def}(n^2)=\frac{n^2}{2}\sum_{a,b,c}\boldsymbol{\mu}_a^T\mathbf{C}^{-1}\boldsymbol{\mu}_b\boldsymbol{\mu}_c^T\mathbf{C}^{-1}\boldsymbol{\mu}_{def}, \tag{B10}$$

where the notation $\Sigma_{(a,b)\times(cd,ef)}$ indicates a sum over all possible permutations of a and b orderings combined with permutations of cd and ef orderings, leading to a total of four terms.

APPENDIX C: JOINT MOMENTS FOR ASYMPTOTIC GAUSSIAN INFERENCE; MULTIVARIATE GAUSSIAN DATA WITH ZERO-MEAN: RANDOM SIGNAL IN NOISE

For this case the mean is zero in Eq. (6). The joint moments required to evaluate the first order bias, as well as the second order error correlation and covariance are then

$$i_{ab}=\frac{n}{2}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b), \tag{C1}$$

$$v_{abc}(n^1)=\frac{n}{3}\sum_{a,b,c}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c) - \frac{n}{4}\sum_{a,b,c}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{ab}\mathbf{C}^{-1}\mathbf{C}_c), \tag{C2}$$

$$v_{a,b,c}(n^1)=\frac{n}{6}\sum_{a,b,c}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c), \tag{C3}$$

$$v_{ab,c}(n^1)=-\frac{n}{2}\sum_{a,b}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c) + \frac{n}{2}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{ab}\mathbf{C}^{-1}\mathbf{C}_c), \tag{C4}$$

$$v_{abcd}(n^1)=-\frac{3n}{8}\sum_{a,b,c,d}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d) + \frac{n}{2}\sum_{a,b,c,d}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{ab}\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d) - \frac{n}{16}\sum_{a,b,c,d}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{ab}\mathbf{C}^{-1}\mathbf{C}_{cd}) - \frac{n}{12}\sum_{a,b,c,d}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{abc}\mathbf{C}^{-1}\mathbf{C}_d), \tag{C5}$$

$$v_{a,b,c,d}(n^2)=\frac{n^2}{32}\sum_{a,b,c,d}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b)\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d), \tag{C6}$$

$$v_{a,b,cd}(n^1)=-\frac{n}{2}\times\sum_{(a,b)\times(c,d)}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d) + \frac{n}{2}\sum_{a,b}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_{cd}), \tag{C7}$$

$$v_{a,b,c,de}(n^2)=-\frac{n^2}{24}\sum_{(a,b,c)\times(d,e)}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_d\mathbf{C}^{-1}\mathbf{C}_e)\times\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c) - \frac{n^2}{8}\sum_{(a,b,c)\times(d,e)}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b)\times\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d\mathbf{C}^{-1}\mathbf{C}_e) + \frac{n^2}{8}\sum_{a,b,c}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b)\times\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_{de}), \tag{C8}$$

$$\begin{aligned}v_{a,b,cd,ef}(n^2)=&\frac{n^2}{8}\sum_{(cd,ef)\times(a,b)\times(c,d)\times(e,f)}\left\{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d)\left[\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_e\mathbf{C}^{-1}\mathbf{C}_f)-\frac{1}{2}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_{ef})\right]\right\}\\&+\frac{n^2}{8}\sum_{(cd,ef)\times(a,b)\times(c,d)\times(e,f)}\left\{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b)\left[\frac{1}{2}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d\mathbf{C}^{-1}\mathbf{C}_e\mathbf{C}^{-1}\mathbf{C}_f)\right.\right.\\&\left.\left.-\frac{1}{2}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d\mathbf{C}^{-1}\mathbf{C}_{ef})+\tfrac{1}{8}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{cd}\mathbf{C}^{-1}\mathbf{C}_{ef})\right]\right\}\\&+\frac{n^2}{8}\sum_{(cd,ef)\times(a,b)\times(c,d)\times(e,f)}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d)\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_e\mathbf{C}^{-1}\mathbf{C}_f)\end{aligned}$$

$$+\frac{n^2}{8}\sum_{(cd,ef)\times(a,b)} \mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_{cd})\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_{ef})$$

$$-\frac{n^2}{8}\sum_{(cd,ef)\times(a,b)\times(c,d)\times(e,f)} \mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_{cd})\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_e\mathbf{C}^{-1}\mathbf{C}_f), \tag{C9}$$

$$v_{a,b,c,def}(n^2)=\frac{n^2}{18}\sum_{(a,b,c)\times(d,e,f)} \mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c)\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_d\mathbf{C}^{-1}\mathbf{C}_e\mathbf{C}^{-1}\mathbf{C}_f)$$

$$-\frac{n^2}{24}\sum_{(a,b,c)\times(d,e,f)} \mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b\mathbf{C}^{-1}\mathbf{C}_c)\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_{de}\mathbf{C}^{-1}\mathbf{C}_f)$$

$$+\frac{n^2}{8}\sum_{(a,b,c)\times(d,e,f)} \{\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_a\mathbf{C}^{-1}\mathbf{C}_b)[\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_d\mathbf{C}^{-1}\mathbf{C}_e\mathbf{C}^{-1}\mathbf{C}_f)$$

$$-\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_{de}\mathbf{C}^{-1}\mathbf{C}_f)+\tfrac{1}{6}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{C}_c\mathbf{C}^{-1}\mathbf{C}_{def})]\}. \tag{C10}$$

The notation $\Sigma_{(a,b,c)\times(d,e,f)}$ indicates summation over all possible permutations of d, e, f and a, b and c, leading to a total of 36 terms.

APPENDIX D: DERIVATION OF THE ASYMPTOTIC EXPANSION OF THE MAXIMUM LIKELIHOOD ESTIMATE

Following Barndoff-Nielsen and Cox,[7] Eq. (1) is first inverted for $(\hat{\theta}-\theta)^r$ to obtain the expansion

$$(\hat{\theta}-\theta)^r=j^{rs}l_s+\tfrac{1}{2}j^{rs}l_{stu}(\hat{\theta}-\theta)^t(\hat{\theta}-\theta)^u+\tfrac{1}{6}j^{rs}l_{stuv}\times(\hat{\theta}-\theta)^t(\hat{\theta}-\theta)^u(\hat{\theta}-\theta)^v+\cdots, \tag{D1}$$

where j^{rs} is the inverse of the observed information matrix $j_{rs}=-l_{rs}$. Iterating this procedure leads to an expression for $(\hat{\theta}-\theta)^r$ that is solely in terms of the derivatives of the likelihood function

$$(\hat{\theta}-\theta)^r=j^{rs}l_s+\tfrac{1}{2}j^{rs}j^{tu}j^{vw}l_{stv}l_ul_w+\tfrac{1}{6}j^{rs}j^{tu}j^{vw}j^{xy}(l_{suyw}+3l_{swp}j^{pq}l_{quy})l_tl_vl_x+\cdots. \tag{D2}$$

The difficulty with this expression is that j^{rs} is not well defined for all likelihood functions and all values of $\hat{\theta}$. This problem is circumvented by expanding j^{rs} in terms of well-defined quantities. First, note that

$$\mathbf{j}=\mathbf{i}\{\mathbf{I}-\mathbf{i}^{-1}(\mathbf{i}-\mathbf{j})\}, \tag{D3}$$

where $\mathbf{I}$ is the identity matrix. The inverse is then

$$\mathbf{j}^{-1}=\{\mathbf{I}-\mathbf{i}^{-1}(\mathbf{i}-\mathbf{j})\}^{-1}\,\mathbf{i}^{-1}. \tag{D4}$$

which can be expanded as

$$\mathbf{j}^{-1}=\mathbf{i}^{-1}+\mathbf{i}^{-1}(\mathbf{i}-\mathbf{j})\mathbf{i}^{-1}+\mathbf{i}^{-1}(\mathbf{i}-\mathbf{j})\mathbf{i}^{-1}(\mathbf{i}-\mathbf{j})\mathbf{i}^{-1}+\cdots, \tag{D5}$$

or equivalently as

$$j^{rs}=i^{rs}+i^{rt}i^{su}H_{tu}+i^{rt}i^{su}i^{vw}H_{tv}H_{uw}+\cdots, \tag{D6}$$

where $H_R=l_R-v_R$ for any set of coordinate indices $R=r_1\ldots r_m$, where, for example, $H_{tu}=l_{tu}-v_{tu}$. Inserting Eq. (D6) into Eq. (D2) then leads to Eq. (2).

[1] C. R. Rao, *Linear Statistical Inference and its Applications* (Wiley, New York, 1966).
[2] A. Battacharya, ''On some analogues of the amount of information and their uses in statistical estimation,'' Sonkhya **8**, 1–14, 201–218, 315–328 (1946).
[3] E. W. Barankin, ''Locally best unbiased estimators,'' Ann. Math. Stat. **20**, 477–501 (1949).
[4] D. Chazan, M. Zakai, and J. Ziv, ''Improved lower bounds on signal parameter estimation,'' IEEE Trans. Inf. Theory **IT-21**, 90–93 (1975).
[5] A. J. Weiss and E. Weinstein, ''Fundamental limitations in passive time delay estimation,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-31**, 472–485 (1983).
[6] R. A. Fisher, *Statistical Methods and Scientific Inference* (Hafner, New York, 1956).
[7] O. E. Barndorff-Nielsen and D. R. Cox, *Inference and Asymptotics* (Chapman & Hall, London, 1994).
[8] H. L. Van Trees, *Detection, Estimation, and Linear Modulation Theory—Part III* (Wiley, New York, 1971).
[9] J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985).
[10] O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt, Eds., *Full Field Inversion Methods in Ocean and Seismic Acoustics* (Kluwer, Dordrecht, 1995).
[11] A. Thode, E. Naftali, Ian Ingram, P. Ratilal, and N. C. Makris, ''Necessary conditions for a maximum likelihood estimate to attain the Cramer-Rao lower bound. Part II: Range and depth localization of a sound source in an ocean waveguide,'' J. Acoust. Soc. Am. (submitted).
[12] A. Zeira and P. Schultheiss, ''Realizable lower bounds for time delay estimation,'' IEEE Trans. Signal Process. **41**, 3102–3113 (1993).
[13] A. B. Baggeroer, ''Barankin bound on the variance of estimates of the parameter of a Gaussian random process,'' M.I.T. Res. Lab., Cambridge, MA, Electron, Quart. Prog. Rep. **92**, 324–333 (1969).
[14] L. R. Shenton and K. O. Bowman, *Maximum Likelihood Estimation in Small Samples* (Charles Griffin, London, 1977).
[15] M. Zanolin, E. Naftali, and N. C. Makris, ''Second order bias of a multivariate Gaussian maximum likelihood estimates with a chain rule for higher moments,'' J. Roy. Statist. Soc. B (submitted).
[16] N. C. Makris, ''Parameter resolution bounds that depend on sample size,'' J. Acoust. Soc. Am. **99**, 2851–2861 (1996). The misstatement in this paper that the distribution for the sample covariance of a CCGR data vector is unknown is corrected in Ref. 17 below. The distribution in question is the complex Wishart distribution.
[17] N. C. Makris, ''The statistics of ocean-acoustic ambient noise,'' in *Sea Surface Sound '97*, edited by T. Leighton (Kluwer Academic, Dordrecht, 1997).
[18] B. R. Kerman, Ed., *Sea Surface Sound, Natural Mechanisms of Surface Generated Noise in the Ocean* (Kluwer, Boston, 1987).
[19] N. C. Makris, ''The effect of saturated transmission scintillation on ocean-acoustic intensity measurements,'' J. Acoust. Soc. Am. **100**, 769–783 (1996).
[20] N. C. Makris, ''A foundation for logarithmic measures of fluctuating intensity in pattern recognition,'' Opt. Lett. **20**, 2012–2014 (1995).
[21] N. C. Makris and P. Ratilal, ''A unified model for reverberation and sub-

merged target scattering in a stratified ocean waveguide,'' J. Acoust. Soc. Am. **109**, 909–941 (2001).
[22] P. Swerling, ''Probability of detection for fluctuating targets,'' Rand Report RM 1217 (1954); reissued in Trans. IRE Prof. Group Inf. Theory **IT-6**, 269–308 (1960).
[23] P. G. Bergmann, ''Intensity fluctuations,'' in *The Physics of Sound in the Sea, Part I: Transmission* (National Defense Research Committee, Washington, DC, 1946).
[24] I. Dyer, ''Statistics of sound propagation in the ocean,'' J. Acoust. Soc. Am. **48**, 337–345 (1970).
[25] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, ''Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem,'' J. Acoust. Soc. Am. **83**, 571–587 (1988).
[26] A. B. Baggeroer, W. A. Kuperman, and P. H. Mikhalevsky, ''An overview of matched field methods in ocean acoustics,'' IEEE J. Ocean Eng. **18**, 401–424 (1993).
[27] Naftali, ''First order bias and second order variance of the Maximum Likelihood Estimator with application to multivariate Gaussian data and time delay and Doppler shift estimation,'' SM Thesis, MIT, Cambridge, MA, 2000.
[28] P. M. Woodward, *Probability and Information Theory, with Application to Radar* (Pergamon, New York, 1964).
[29] N. C. Makris, C. S. Chia, and L. T. Fialkowski, ''The bi-azimuthal scattering distribution of an abyssal hill,'' J. Acoust. Soc. Am. **106**, 2491–2512 (1999).