Parameter Estimation of 2-D Random Amplitude Polynomial-Phase Signals

Joseph M. Francos and Benjamin Friedlander, *Fellow, IEEE*

***Abstract*—Phase information has fundamental importance in many two-dimensional (2-D) signal processing problems. In this paper, we consider 2-D signals with random amplitude and a continuous deterministic phase. The signal is represented by a random amplitude polynomial-phase model. A computationally efficient estimation algorithm for the signal parameters is presented. The algorithm is based on the properties of the mean phase differencing operator, which is introduced and analyzed. Assuming that the signal is observed in additive white Gaussian noise and that the amplitude field is Gaussian as well, we derive the Cramér-Rao lower bound (CRB) on the error variance in jointly estimating the model parameters. The performance of the algorithm in the presence of additive white Gaussian noise is illustrated by numerical examples and compared with the CRB.**

I. INTRODUCTION

PHASE information has fundamental importance in many one- and two-dimensional (1-D and 2-D) signal processing problems. When dealing with 2-D signals, estimates of the phase are required in different applications such as 2-D homomorphic signal processing, magnetic resonance imaging (MRI), [1]–[3], optical imaging, [4], and interferometric synthetic aperture radar (INSAR), [5], [6]. In processing nonstationary 1-D signals, as well as in the case of nonhomogeneous multidimensional signals, the phase contains useful information. In 1-D signals, the first derivative of the phase is the instantaneous frequency of the signal, whereas for multidimensional data, the partial derivatives of the phase along each of the spatial axes provide the local spatial frequency of the analyzed field.

Recently, an algorithm for estimating the shape of a 3-D object, based on a single image of its textured surface, has been presented [13]. The algorithm employs a nonparametric estimation method to compute the local phase function of the object image. The local phase information is then employed for calculating the normal to the object surface.

In SAR imaging, the amplitude of the received complex valued 2-D image is proportional to the backscattering of the illuminated points. In interferometric SAR, two images $I_1(n,m)$ and $I_2(n,m)$ are obtained from two different antennas illuminating the same target point. Taking the conjugated product $I_1(n,m)I_2^*(n,m)$ of these two images, the interferometric SAR (INSAR) signal is obtained. The phase of this 2-D INSAR signal is proportional to the elevation of the scattering point on the ground. Hence, ground elevations and terrain maps can be produced from the INSAR data [5], [6]. A critical consideration in producing the three-dimensional (3-D) terrain maps is the need to perform 2-D phase unwrapping of the observed signal phase to enable a meaningful interpretation of the data. Ideally, in the absence of noise and phase aliasing, we could unwrap the phase function by following an integration path and adding multiples of 2π to the phase whenever a sudden drop from π to $-\pi$ occurs. To ensure that no phase-aliasing occurs, the original scene must be properly sampled so that phase differences between two adjacent samples are smaller than π rad. This requirement cannot be generally satisfied, and hence, in the presence of noise and phase aliasing, this simple phase unwrapping method is inadequate.

In this paper, we address the problem of estimating the parameters of such 2-D signals. More specifically, we consider here 2-D signals with random amplitude and a continuous phase function. In these signals, the phase is the information of interest, whereas the random amplitude is a multiplicative noise that highly complicates the phase estimation. Since continuous functions can be approximated by polynomials, a natural choice for modeling the signal phase is by a 2-D polynomial function of the coordinates. Having estimated the phase of the signal, it is a straightforward task to obtain estimates of its local spatial frequencies as well. In this paper, we address separately the cases where the random amplitude field is of a nonzero mean and the case where the amplitude field is a zero mean field. A good example of a positive amplitude field is that of the INSAR image. Assuming the amplitude field of each of the SAR images $I_1(n,m)$ and $I_2(n,m)$ has a Rayleigh probability density function, the amplitude of $I_1(n,m)I_2^*(n,m)$ has an exponential probability density function.

We will derive a computationally efficient algorithm for estimating the parameters of 2-D random-amplitude polynomial phase signals. Such an algorithm can serve as a basic building block in processing INSAR data. The proposed algorithm is an extension of the 1-D parameter estimation algorithms proposed in [7] and [10] and of the algorithm for estimating the parameters of 2-D complex valued, constant amplitude, polynomial phase signals [8]. The algorithm in [7] uses the high-order ambiguity function [18] to estimate the parameters of 1-D complex valued, constant amplitude, polynomial-phase signals. This algorithm is adapted in [10] to

Manuscript received September 12, 1996; revised December 22, 1998. This work was supported by the Office of Naval Research under Contract N00014-95-1-0912. The associate editor coordinating the review of this paper and approving it for publication was Dr. Petar M. Djuric.

J. M. Francos is with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, Israel.

B. Friedlander is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA.

Publisher Item Identifier S 1053-587X(99)04657-7.

estimate the parameters of 1-D random amplitude polynomial-phase signals. The algorithm derived here is based on the properties of a 2-D mean phase difference operator, which is defined in the next section.

The paper is organized as follows. In Section II, we define the parametric model of the observed signal, define the 2-D mean phase difference operator, and present some properties of the operator. In Section III we present a parameter estimation algorithm based on the 2-D mean phase difference operator and its properties. In the first part of this section, we present the algorithm for the case of a nonzero mean random amplitude field, and in the second part, we present a modification for the case of a zero mean amplitude field. These algorithms require knowledge of the observed signal moments, which are not available to us. Therefore, in Section IV, we describe a method for applying the mean phase difference operator when we are given a single observed realization of the field. In Section V, we address the problem of estimating the parameters of the random-amplitude polynomial phase signal in the presence of observation noise. In Section VI we derive the exact Cramér-Rao lower bound (CRB) on the accuracy of estimating the model parameters for a polynomial phase signal with Gaussian random amplitude. This derivation is then specialized for the case where the observations are known to be at a high signal-to-noise ratio (SNR). In Section VI we illustrate the operation of the proposed algorithm in the presence of observation noise using some numerical examples and Monte Carlo simulations.

II. Phase Difference Operator

In this section, we define the phase difference operator and present some of its basic properties. We start with a description of the type of signal for which the operator was designed.

A. Signal Model

Let $\{y(n,m)\}$ be a discrete 2-D random field consisting of the sum of a random amplitude polynomial-phase signal and additive white Gaussian noise. More specifically

$$y(n,m) = v(n,m) + u(n,m), \quad n = 0,1,\ldots,N-1, \quad m = 0,1,\ldots,M-1 \tag{1}$$

where

$$v(n,m) = w(n,m)\cdot \exp\{j\phi_{S+1}(n,m)\} \tag{2}$$

$$\phi_{S+1}(n,m) = \sum_{(k,\ell)\in I} c(k,\ell) n^k m^\ell \tag{3}$$

and $I = \{0 \le k,\ell \text{ and } 0 \le k+\ell \le S+1\}$. We call $\phi_{S+1}(n,m)$ a 2-D polynomial of *total-degree* $S+1$ [8]. Intuitively, we might think of the phase polynomial $\phi_S(n,m)$ as if it has S "layers" since increasing S by one adds a layer of additional $S+2$ parameters to the phase model. To further illustrate the definition, we depict, in Fig. 1, a triangular support of total-degree 4.

The amplitude field $\{w(n,m)\}$ is an ergodic, real-valued, *strict sense homogeneous* random field. The observation noise $u(n,m)$ is assumed to be complex valued, zero mean, circular white Gaussian noise. It is assumed to be independent of the amplitude field $\{w(n,m)\}$. In this section, in order to simplify the presentation, we discuss the case in which there is no observation noise. Hence, $y(n,m) = v(n,m)$.

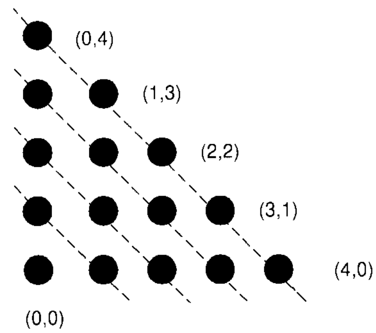


Fig. 1. Triangular support of total-degree 4. Diagonal lines indicate layers 1 through 4.

B. Mean Phase Differencing Operators

Next, we define the two polynomial phase difference operators, which we denote by PD_n and PD_m. We start with a brief heuristic explanation of the idea behind the operators.

Consider the signal given by (2) and (3), and assume for the moment that m and n are continuous variables and that $w(n,m) = A$, where A is some positive deterministic constant. Differentiating the phase of the observed signal P times along the m axis and $S-P$ times along the n axis (in any order as long as the total number of differentiation operations in both axes is S) results in a 2-D complex exponential signal. It can be shown that the spatial frequency (ω,ν) of this complex exponential is a function of two of the coefficients of the highest layer $S+1$ of the phase polynomial parameters and other known quantities. By estimating the frequency of the complex exponential (using standard frequency estimation techniques), we obtain estimates of two of the coefficients of the highest layer of the phase polynomial model. Repeating this procedure for all $0 \le P \le S$, all the coefficients of the highest layer of the phase model are estimated.

Having completed the estimation of the phase parameters in the highest layer, their contribution to the signal phase can be eliminated, thus resulting in a polynomial phase signal of total-degree S. By repeating this entire process for all the layers in the phase model, all the phase parameters are estimated. The details of how that works will be presented later.

Since, in our problem, the variables n and m are discrete, phase differentiating will be replaced by phase differencing. We next define the basic phase differencing operators.

Definition 1 [8]: Let τ_m and τ_n be some strictly positive integers. Define

$$\mathrm{PD}_{m^{(0)}}[v(n,m)] \triangleq v(n,m), \quad n = 0,1,\ldots,N-1, \quad m = 0,1,\ldots,M-1 \tag{4}$$

and in general

$$\mathrm{PD}_{m^{(q)}}[v(n,m)] \triangleq \mathrm{PD}_{m^{(q-1)}}[v(n,m)](\mathrm{PD}_{m^{(q-1)}}[v(n,m+\tau_m)])^* \tag{5}$$

where the resulting 2-D signal $\mathrm{PD}_{m^{(q)}}[v(n,m)]$ exists for $n=0,1,\ldots,N-1$, $m=0,1,\ldots,M-1-q\tau_m$. Similarly

$$\mathrm{PD}_{n^{(0)}}[v(n,m)] \triangleq v(n,m) \quad n=0,1,\ldots,N-1, \quad m=0,1,\ldots,M-1 \tag{6}$$

and

$$\begin{aligned}&\mathrm{PD}_{n^{(p)}}[v(n,m)]\\ &\quad\triangleq \mathrm{PD}_{n^{(p-1)}}[v(n,m)](\mathrm{PD}_{n^{(p-1)}}[v(n+\tau_n,m)])^*\\ &n=0,1,\ldots,N-1-p\tau_n, \quad m=0,1,\ldots,M-1.\end{aligned} \tag{7}$$

Definition 2: Let p and q be some positive integers. Define

$$\overline{\mathrm{PD}}_{m^{(q)}}[v(n,m)] \triangleq E\{\mathrm{PD}_{m^{(q)}}[v(n,m)]\} \tag{8}$$

$$\overline{\mathrm{PD}}_{n^{(p)}}[v(n,m)] \triangleq E\{\mathrm{PD}_{n^{(p)}}[v(n,m)]\}. \tag{9}$$

We shall call these operators the *mean phase difference* (MPD) operators.

The operators are called "phase differencing operators" since they perform an operation that is equivalent to phase differentiation of a continuous parameter 2-D phase [8]. Later in this section, we provide an alternative representation and interpretation of the operators $\overline{\mathrm{PD}}_{m^{(q)}}[\cdot]$ and $\overline{\mathrm{PD}}_{n^{(p)}}[\cdot]$. Note that applying any of the operators $\overline{\mathrm{PD}}_{m^{(1)}}[\cdot]$ or $\overline{\mathrm{PD}}_{n^{(1)}}[\cdot]$ to a 2-D random amplitude polynomial phase signal of total-degree $S+1$ results in a *constant amplitude* (in m and m) 2-D polynomial phase signal of total-degree S.

Some of the properties of the MPD operator are more easily proven using the properties of the ∇_m and ∇_n difference operators, which were introduced in [9]. Next, we repeat the definitions and briefly summarize the main properties of these operators.

Definition 3: Let τ_m and τ_n be some strictly positive integers. The ∇_m*-difference operator* of a 2-D function $\phi(n,m)$ is a linear operator defined by

$$\nabla_m[\phi(n,m)] = \phi(n,m) - \phi(n,m+\tau_m) \tag{10}$$

i.e., ∇_m is a difference operator along the m axis. Similarly, the ∇_n*-difference operator* is defined by $\nabla_n[\phi(n,m)] = \phi(n,m) - \phi(n+\tau_n,m)$.

It is straightforward to show, using the definitions and the linearity of the operators, that the difference operations are commutative, i.e., $\nabla_n[\nabla_m[\phi(n,m)]] = \nabla_m[\nabla_n[\phi(n,m)]]$. Hence, applying P times the ∇_n difference operator and $S-P$ times the ∇_m difference operator to $\phi(n,m)$ yields a unique result, irrespective of the order in which the operators were applied to $\phi(n,m)$. In the following, we denote the resulting signal by $\nabla_{n^{(P)},m^{(S-P)}}[\phi(n,m)]$. Let $\phi_{S+1}(n,m)$ be a 2-D polynomial of total-degree $S+1$. Then, it is shown in [9] that

$$\nabla_{n^{(P)},m^{(S-P)}}[\phi_{S+1}(n,m)] = \omega_S n + \nu_S m + \gamma_S(\tau_n,\tau_m) \tag{11}$$

where

$$\omega_S = (-1)^S c(P+1,S-P)(P+1)!(S-P)!\tau_n^P\tau_m^{S-P} \tag{12}$$

$$\nu_S = (-1)^S c(P,S+1-P)P!(S+1-P)!\tau_n^P\tau_m^{S-P} \tag{13}$$

and $\gamma_S(\tau_n,\tau_m)$ is neither a function of m nor n. It was also shown that applying P times, in arbitrary order, the operator ∇_n and $K-P$ times the operator ∇_m to $\phi(n,m)$ yields

$$\begin{aligned}\nabla_{n^{(P)},m^{(K-P)}}[\phi(n,m)] = &\sum_{q=0}^{K-P}\sum_{p=0}^{P}(-1)^{p+q}\binom{P}{p}\\ &\times\binom{K-P}{q}\phi(n+p\tau_n,m+q\tau_m).\end{aligned} \tag{14}$$

C. Alternative Representation of the $\overline{PD}_n$ and $\overline{PD}_m$ Operators

Based on the properties of the ∇_m and ∇_n difference operators and Definition 1, it can be easily verified that applying P times the phase difference operator $\mathrm{PD}_{n^{(1)}}$ and $S-P$ times the phase difference operator $\mathrm{PD}_{m^{(1)}}$ to the signal $v(n,m)$ yields a unique result, irrespective of the order in which the operators were applied to $v(n,m)$. In the following, we denote the resulting signal by $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(n,m)]$.

Define

$$\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)] = E\{\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(n,m)]\}. \tag{15}$$

Lemma 1:

$$\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)] = E\left\{\prod_{q=0}^{S-P}\left\{\prod_{p=0}^{P}[v^{(*^{(p+q)})}(n+p\tau_n,m+q\tau_m)]^{\binom{P}{p}}\right\}^{\binom{S-P}{q}}\right\} \tag{16}$$

where we define

$$\begin{aligned}&v^{(*^{(p+q)})}(n+p\tau_n,m+q\tau_m)\\ &\quad=\begin{cases}v(n+p\tau_n,m+q\tau_m), & p+q \text{ even}\\ v^*(n+p\tau_n,m+q\tau_m), & p+q \text{ odd}.\end{cases}\end{aligned} \tag{17}$$

Proof: The proof is an immediate extension of Lemma 2 in [9]. □

Theorem 1: Let $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(n,m)]$ be the 2-D signal obtained by successively applying, in some arbitrary sequence, P times the operator $\mathrm{PD}_{n^{(1)}}[\cdot]$ and $S-P$ times the operator $\mathrm{PD}_{m^{(1)}}[\cdot]$ to the signal (2). Then, the signal $\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]$ is a 2-D exponential given by

$$\begin{aligned}&\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]\\ &\quad=\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)]\\ &\quad\quad\cdot\exp\{j[\omega_S n+\nu_S m+\gamma_S(\tau_n,\tau_m)]\}\\ &\quad\quad n=0,1,\ldots,N-1-P\tau_n\\ &\quad\quad m=0,1,\ldots,M-1-(S-P)\tau_m\end{aligned} \tag{18}$$

where

$$\omega_S = (-1)^S c(P+1, S-P)(P+1)!(S-P)!\tau_n^P \tau_m^{S-P} \tag{19}$$

$$\nu_S = (-1)^S c(P, S+1-P)P!(S+1-P)!\tau_n^P \tau_m^{S-P} \tag{20}$$

while both $\overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)]$ and $\gamma_S(\tau_n, \tau_m)$ are neither functions of m nor n.

Proof: Consider the 2-D signal

$$\begin{aligned}
&\overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)] \cdot \exp\{j[\omega_S n + \nu_S m + \gamma_S(\tau_n, \tau_m)]\} \\
&= \overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)] \\
&\quad \cdot \exp\{j\nabla_{n^{(P)},m^{(S-P)}}[\phi_{S+1}(n,m)]\} \\
&= \overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)] \cdot \exp\Bigg\{ j \sum_{q=0}^{S-P} \sum_{p=0}^{P} (-1)^{p+q} \\
&\quad \times \binom{P}{p}\binom{S-P}{q} \phi_{S+1}(n+p\tau_n, m+q\tau_m) \Bigg\} \\
&= \overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)] \prod_{q=0}^{S-P} \prod_{p=0}^{P} \exp\Bigg\{ j(-1)^{p+q} \\
&\quad \times \binom{P}{p}\binom{S-P}{q} \phi_{S+1}(n+p\tau_n, m+q\tau_m) \Bigg\} \\
&= E\Bigg\{ \prod_{q=0}^{S-P} \Bigg\{ \prod_{p=0}^{P} [w^{(*^{(p+q)})}(n+p\tau_n, m \\
&\quad + q\tau_m)]^{\binom{P}{p}} \Bigg\}^{\binom{S-P}{q}} \Bigg\} \\
&\quad \cdot \prod_{q=0}^{S-P} \Bigg\{ \prod_{p=0}^{P} \{[\exp\{j\phi_{S+1}(n+p\tau_n, m \\
&\quad + q\tau_m)\}]^{(-1)^{p+q}}\}^{\binom{P}{p}} \Bigg\}^{\binom{S-P}{q}} \\
&= E\Bigg\{ \prod_{q=0}^{S-P} \Bigg\{ \prod_{p=0}^{P} \{[v^{(*^{(p+q)})}(n+p\tau_n, m \\
&\quad + q\tau_m)]\}^{\binom{P}{p}} \Bigg\}^{\binom{S-P}{q}} \Bigg\} \\
&= \overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]
\end{aligned} \tag{21}$$

where the first equality is due to (11), the second equality is due to (14), and the last equality is due to Lemma 1. Since $\{w(n,n)\}$ is a strict sense homogeneous random field, its statistics are invariant to a shift of the origin. Hence, its moments of any order are independent of n and m but, rather, are functions of coordinate differences. □

III. Parameter Estimation Algorithm

A. Estimation Procedure for a Nonzero Mean Amplitude Field

Consider the signal given by (2) and (3), where S is a non-negative integer, which is assumed initially to be known. We now present an algorithm for sequentially estimating the parameters $\{c(k,\ell) \mid 0 \leq k, \ell; 0 \leq k+\ell \leq S+1\}$ of the 2-D random amplitude polynomial phase signal, where it is *a priori* known that the amplitude field has a nonzero mean.

Theorem 1 implies that applying P times the operator $\text{PD}_{n^{(1)}}$ and $S-P$ times the operator $\text{PD}_{m^{(1)}}$ to $v(n,m)$, followed by taking the expectation of the resulting signal, we obtain the 2-D exponential (18). We can thus reduce any 2-D nonhomogeneous, random-amplitude polynomial-phase signal, whose phase is of total-degree $S+1$, to a 2-D sinusoidal signal whose frequency is (ω_S, ν_S). Hence, estimating (ω_S, ν_S) using any standard frequency estimation technique results in an estimate of $c(P+1, S-P)$ and $c(P, S+1-P)$. In this paper, we estimate the frequency of the exponential using a search for the maximum of the absolute value of the 2-D discrete Fourier transform (2-D DFT) of the signal.

Note from (19) and (20) that the phase coefficients can be estimated unambiguously (i.e., with no aliasing) as long as

$$|c(P+1, S-P)| \leq \frac{\pi}{(P+1)!(S-P)!\tau_n^P \tau_m^{S-P}} \tag{22}$$

and similarly for $c(P, S+1-P)$. However, since a parametric model is fitted to the observed signal, the phase function itself can be sampled *under* the Nyquist rate because the phase estimation is not performed through a waveform-based procedure. Therefore, phase differences between adjacent samples may be greater than π rad without affecting the ability of the algorithm to estimate the phase parameters, as long as the constraint (22) is satisfied. In other words, the proposed phase-estimation algorithm can perform very well when phase aliasing due to low sampling and noise are present. This point is further illustrated in Section VII.

Thus, having estimated ω_S and ν_S in (19) and (20), we have

$$\hat{c}(P+1, S-P) = \frac{\hat{\omega}_S}{(-1)^S(P+1)!(S-P)!\tau_n^P \tau_m^{S-P}} \tag{23}$$

and

$$\hat{c}(P, S+1-P) = \frac{\hat{\nu}_S}{(-1)^S P!(S+1-P)!\tau_n^P \tau_m^{S-P}} \tag{24}$$

which constitutes an estimate of two of the parameters of the highest order layer $S+1$ of the phase model parameters (i.e., those $c(k,\ell)$'s for which $0 \leq k, \ell : k+\ell = S+1$).

Recall, however, that the $S+1$ layer has $S+2$ parameters that need to be estimated. This can be achieved by repeating the procedure described above, assuming some arbitrary P, for *all* P such that $0 \leq P \leq S$. Note that this procedure results in repeated estimation of some of the phase parameters. Let

$$\psi_Q(n,m) = \sum_{k=0}^{Q} \hat{c}(k, Q-k) n^k m^{Q-k} \tag{25}$$

denote the estimated Qth layer of the phase function.

Multiplying $v(n,m)$ by $\exp\{-j\psi_{S+1}(n,m)\}$ results in a new random amplitude polynomial phase signal whose total degree is S. By applying to the resulting signal a procedure similar to the one used to estimate the parameters of the $S+1$ layer, we obtain an estimate of the $S+1$ parameters in the S layer. Multiplying the 2-D random-amplitude polynomial phase signal of total degree S, which was obtained in the previous step, by $\exp\{-j\psi_S(n,m)\}$, we obtain a new random-amplitude polynomial-phase signal whose total degree is $S-1$.

In general, let $v^{(s+1)}(n,m)$ denote the 2-D signal where $s+1$ denotes the *current* total degree of its phase polynomial. The phase parameters are sequentially estimated, layer after layer, for all $s=S,\ldots,0$. For each layer, the algorithm is a two-stage procedure. In the first stage, the parameters of layer $s+1$ are estimated by finding, for all $0\leq P\leq s$, the maxima of the absolute value of the DFT of $\overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[v^{(s+1)}(n,m)]$. In the second stage, the already-reduced order 2-D random-amplitude polynomial phase signal is multiplied by $\exp\{-j\psi_{s+1}(n,m)\}$.

Using this procedure, we obtain estimates for all the phase parameters except $c(0,0)$. The signal resulting from this processing is denoted by $v^{(0)}(n,m)$. If the amplitude field $\{w(n,m)\}$ is known to be positive for all (n,m) (e.g., the amplitude field is exponentially distributed) then, by taking the average of the imaginary part of the logarithm of $v^{(0)}(n,m)$, we obtain an estimate for $c(0,0)$. In general, the amplitude field can assume both negative and positive values. Hence, $c(0,0)$ can only be estimated up to a magnitude π factor. More specifically, we assume that $c(0,0)\in[0,\pi)$. Thus, let

$$\tilde{\phi}^{(0)}(n,m)=\begin{cases}\text{Im}\big\{\log\big(v^{(0)}(n,m)\big)\big\}, & \text{Im}\big\{\log\big(v^{(0)}(n,m)\big)\big\}\geq 0\\ \text{Im}\big\{\log\big(v^{(0)}(n,m)\big)\big\}+\pi, & \text{Im}\big\{\log\big(v^{(0)}(n,m)\big)\big\}<0.\end{cases} \quad (26)$$

Taking the average of $\tilde{\phi}^{(0)}(n,m)$, we obtain an estimate for $c(0,0)$. We have thus completed the estimation of all the coefficients of the 2-D phase polynomial of total degree $S+1$. It should be noted that if the amplitude is positive, the estimation algorithm of the phase parameters is *identical* to the algorithm derived in [8] for *constant* amplitude polynomial phase signals, even though here we are dealing with *random* amplitudes. The estimation problem when $\{w(n,m)\}$ is a zero mean random field is discussed in Section III-B.

Once the phase parameters were estimated, the random amplitude of the polynomial phase signal is obtained by multiplying the observed signal by $e^{-j\hat{\phi}(n,m)}$, where $\hat{\phi}(n,m)$ is the estimated phase. Since $\{w(n,m)\}$ is a homogeneous random field, its parameters can be estimated using any standard algorithm (see, e.g., [14] for the case where $\{w(n,m)\}$ is an autoregressive field and [16] for the case where $\{w(n,m)\}$ is a moving-average field).

Finally, we note that for a 2-D random amplitude polynomial phase signal $v(n,m)$ of total degree S, $\overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]$ is neither a function of n nor m. As we show in Section VI, the CRB on the phase parameters is independent of their values. These two properties allow for relatively simple order estimation in cases where the polynomial total degree S is unknown, but the amplitude field is *a priori* known to be Gaussian. Assume an arbitrary upper bound on the total degree S. In the presence of observation noise, we decide that $c(k,\ell)=0$ whenever $|\hat{c}(k,\ell)|$ is not considerably higher than $\{\text{CRB}[c(k,\ell)]\}^{\frac{1}{2}}$.

B. Estimating the Parameters of Signals with a Zero Mean Amplitude

Adopting the approach described above for the case of signals with zero mean amplitude yields estimates of all the phase parameters except $c(0,0)$ and the first layer parameters $c(0,1)$ and $c(1,0)$. To see this, consider a zero-mean random amplitude polynomial phase signal whose total degree is 1, i.e.,

$$v(n,m)=w(n,m)\cdot\exp\{j[c(1,0)n+c(0,1)m+c(0,0)]\}. \quad (27)$$

Since $\{w(n,m)\}$ is a zero mean random field, applying to this random-amplitude exponential signal the MPD operator $\overline{\text{PD}}_{n^{(0)},m^{(0)}}[\cdot]$ results in a zero signal for all n and m. Hence, the algorithm proposed for estimating the parameters of higher layers is useless in the case where $k+\ell=1$. We must therefore resort to an alternative algorithm for estimating these parameters. Next, we redefine the operator $\overline{\text{PD}}_{n^{(0)},m^{(0)}}$ to avoid this problem.

Definition 4: Let τ_m and τ_n be some strictly positive integers. Define

$$\overline{\text{PD}}_{n^{(0)},m^{(0)}}[v(n,m)]\triangleq E[v(n,m)v(n+\tau_n,m+\tau_m)]. \quad (28)$$

For the case in which $v(n,m)$ is a random-amplitude polynomial phase signal of total degree 1, we have that

$$\begin{aligned}\overline{\text{PD}}_{n^{(0)},m^{(0)}}&[v(n,m)]\\ &=E[w(n,m)w(n+\tau_n,m+\tau_m)]\cdot\exp\{j[2c(1,0)n\\ &\quad+2c(0,1)m+(2c(0,0)+c(1,0)\tau_n+c(0,1)\tau_m)]\}.\end{aligned} \quad (29)$$

Since $\{w(n,m)\}$ is strict sense homogeneous, $E[w(n,m)\,w(n+\tau_n,m+\tau_m)]$ is neither a function of n nor m. Hence, $\overline{\text{PD}}_{n^{(0)},m^{(0)}}[v(n,m)]$ is a constant amplitude exponential whose frequency is $(2c(1,0),2c(0,1))$. The exponential frequency can be estimated using any standard frequency estimation technique. Finally, $c(0,0)$ is estimated using the procedure that was described in Section III-A for a random amplitude field that is not necessarily positive. The algorithm for the case of a zero-mean amplitude is summarized in Table I.

IV. Estimation of the Observed Signal Moments

The algorithms presented in Section III are formulated in terms of high-order moments of $v(n,m)$. In this section, we address the problem of estimating the moments of this nonhomogeneous field when only a finite single observed realization of the field is available. Clearly, since the field is nonhomogeneous, it is also nonergodic. Hence, a straightforward replacement of ensemble averages by spatial averages is impossible.

More specifically, we are interested in estimating $\overline{\text{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)],\ n=0,1,\ldots,N-1-P\tau_n,\ m=0,1,$

TABLE I
ESTIMATION ALGORITHM FOR A ZERO-MEAN AMPLITUDE FIELD

Let $S+1$ denote the total-degree of the observed signal phase.

$s=S,\ v^{(s+1)}(n,m)=v(n,m),\ n=0,\ldots,N-1,\ m=0,\ldots,M-1.$

While $s \geq 1$ ($s+1$ is the layer index)

for $P=0,\ldots,s$ (find all the parameters of the $s+1$ layer)

$(\hat{\omega}_s,\hat{\nu}_s)=\underset{(\omega,\nu)}{\operatorname{argmax}}\left|\mathrm{DFT}\left(\overline{\mathrm{PD}}_{n^{(P)},m^{(s-P)}}[v^{(s+1)}(n,m)]\right)\right|$

$\hat{c}(P+1,s-P)=\frac{\hat{\omega}_s}{(-1)^s(P+1)!(s-P)!\tau_n^P\tau_m^{s-P}}$

$\hat{c}(P,s+1-P)=\frac{\hat{\nu}_s}{(-1)^s P!(s+1-P)!\tau_n^P\tau_m^{s-P}}$

end

$v^{(s)}(n,m)=v^{(s+1)}(n,m)\cdot\exp\{-j\sum_{\{k+\ell=s+1\}}\hat{c}(k,\ell)n^k m^\ell\}$

s=s-1

end

$(\hat{\omega}_1,\hat{\nu}_1)=\underset{(\omega,\nu)}{\operatorname{argmax}}\left|\mathrm{DFT}\left(\overline{\mathrm{PD}}_{n^{(0)},m^{(0)}}[v^{(1)}(n,m)]\right)\right|$

$\hat{c}(1,0)=\frac{\hat{\omega}_1}{2}$

$\hat{c}(0,1)=\frac{\hat{\nu}_1}{2}$

$v^{(0)}(n,m)=v^{(1)}(n,m)\cdot\exp\{-j\sum_{\{k+\ell=1\}}\hat{c}(k,\ell)n^k m^\ell\}$

$\hat{c}(0,0)=\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\tilde{\phi}^{(0)}(n,m)$

$\hat{w}(n,m)=v^{(0)}(n,m)\cdot\exp\{-j\hat{c}(0,0)\}$

$\ldots,M-1-(S-P)\tau_m$. From Theorem 1, we have that

$$\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]=E\{\mathrm{PD}_{n^{(P)},m^{(S-P)}}[w(n,m)]\}\cdot\exp\{j[\omega_S n+\nu_S m+\gamma_S]\}. \tag{30}$$

The term $E\{\mathrm{PD}_{n^{(P)},m^{(S-P)}}[w(n,m)]\}$ of (30) is a high-order moment of a strict sense homogeneous and ergodic random field. Therefore, it can be consistently estimated by replacing the ensemble average with sample average. Hence

$$\begin{aligned}\widehat{\overline{\mathrm{PD}}}_{n^{(P)},m^{(S-P)}}[w(n,m)] &= \frac{1}{(N-P\tau_n)(M-(S-P)\tau_m)} \times\sum_{k=0}^{N-1-P\tau_n}\sum_{\ell=0}^{M-1-(S-P)\tau_m}\mathrm{PD}_{n^{(P)},m^{(S-P)}}[w(k,\ell)]\\ &= \frac{1}{(N-P\tau_n)(M-(S-P)\tau_m)} \times\sum_{k=0}^{N-1-P\tau_n}\sum_{\ell=0}^{M-1-(S-P)\tau_m}\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(k,\ell)]\\ &\quad\cdot\exp\{-j[\omega_S k+\nu_S\ell+\gamma_S]\}\end{aligned} \tag{31}$$

where the last equality can be verified by following the steps of the proof of Theorem 1 with the $\overline{\mathrm{PD}}$ operator replaced by the PD operator. Thus, an estimate of $\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]$ is given by

$$\begin{aligned}\widehat{\overline{\mathrm{PD}}}_{n^{(P)},m^{(S-P)}}[v(n,m)] &= \frac{1}{(N-P\tau_n)(M-(S-P)\tau_m)}\\ &\quad\cdot\left[\sum_{k=0}^{N-1-P\tau_n}\sum_{\ell=0}^{M-1-(S-P)\tau_m}\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(k,\ell)]\times\exp\{-j[\omega_S k+\nu_S\ell+\gamma_S]\}\right]\\ &\quad\cdot\exp\{j[\omega_S n+\nu_S m+\gamma_S]\}\\ &= \frac{1}{(N-P\tau_n)(M-(S-P)\tau_m)}\\ &\quad\cdot\left[\sum_{k=0}^{N-1-P\tau_n}\sum_{\ell=0}^{M-1-(S-P)\tau_m}\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(k,\ell)]\times\exp\{-j[\omega_S k+\nu_S\ell]\}\right]\\ &\quad\cdot\exp\{j[\omega_S n+\nu_S m]\}.\end{aligned} \tag{32}$$

Note that (32) is the 2-D Fourier series expansion of $\widehat{\overline{\mathrm{PD}}}_{n^{(P)},m^{(S-P)}}[v(n,m)]$. The series has a single term. The coefficient of this term is the 2-D Fourier transform applied to the signal $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(n,m)]$, evaluated at some frequency (ω_S,ν_S), and scaled by a constant. Since ω_S and ν_S are unknown, this expression has to be evaluated for all (ω_S,ν_S).

Thus, in the estimation algorithm, we replace the MPD operator $\overline{\mathrm{PD}}$, which is using ensemble moments, with the $\widehat{\overline{\mathrm{PD}}}$ operator, which is using sample moments. More specifically, the step in which we evaluate $(\hat{\omega}_s,\hat{\nu}_s)$ is now replaced by

$$(\hat{\omega}_s,\hat{\nu}_s)=\underset{(\omega,\nu)}{\operatorname{argmax}}\left|\mathrm{DFT}\left(\widehat{\overline{\mathrm{PD}}}_{n^{(P)},m^{(s-P)}}\left[v^{(s+1)}(n,m)\right]\right)\right|. \tag{33}$$

Using the definition of the DFT, it can be verified that the maximization in (33) is achieved when the absolute value of the single coefficient in the Fourier series expansion of $\widehat{\overline{\mathrm{PD}}}_{n^{(P)},m^{(s-P)}}[v^{(s+1)}(n,m)]$ is maximized. In other words, evaluating the Fourier transform of $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v^{(s+1)}(n,m)]$ for all (ω,ν) and setting

$$(\hat{\omega}_s,\hat{\nu}_s)=\underset{(\omega,\nu)}{\operatorname{argmax}}\left|\mathrm{DFT}\left(\mathrm{PD}_{n^{(P)},m^{(s-P)}}\left[v^{(s+1)}(n,m)\right]\right)\right| \tag{34}$$

is equivalent to estimating $(\hat{\omega}_s,\hat{\nu}_s)$ using (33). In conclusion, when only a single realization of the field is observed, $\hat{\omega}_s$ and $\hat{\nu}_s$ are estimated by finding the maxima of the DFT of $\mathrm{PD}_{n^{(P)},m^{(s-P)}}[v^{(s+1)}(n,m)]$. Substitution of the estimates into (23) and (24) provides the desired estimates of the polynomial phase parameters.

Using the derivation of the $\widehat{\overline{\mathrm{PD}}}$ estimator in (31) and (32), it is clear that since $\{w(n,m)\}$ is a strict-sense homogeneous and ergodic random field

$$\lim_{N,M\to\infty} \widehat{\overline{\mathrm{PD}}}_{n^{(P)},m^{(s-P)}}\left[v^{(s+1)}(n,m)\right] = \overline{\mathrm{PD}}_{n^{(P)},m^{(s-P)}}\left[v^{(s+1)}(n,m)\right] \quad (35)$$

which is (see Theorem 1) a *constant* amplitude exponential with the correct frequency (ω_s, ν_s). In other words, the ergodicity of $\{w(n,m)\}$ guarantees that as $N \to \infty$ and $M \to \infty$, $(\hat{\omega}_s, \hat{\nu}_s) = (\omega_s, \nu_s)$. Note, however, that when the dimensions of the observed field are finite, in order for $(\hat{\omega}_s, \hat{\nu}_s)$ to be correctly estimated, the Fourier series coefficient in (32) has to be nonzero and slowly varying relative to (ω_s, ν_s). Clearly, these requirements are satisfied when the mean component of the amplitude signal is larger than its standard deviation. Furthermore, the foregoing discussion implies that even in cases where the amplitude field $w(n,m)$ is nonergodic, but the coefficient in (32) is nonzero and slowly varying as a function of frequency, the phase parameters are correctly estimated, despite the violation of the ergodicity assumption.

An alternative view point of the motivation in adopting the statistic $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v^{(s+1)}(n,m)]$ in (34) is the following: From the derivation of the estimator and the proof of Theorem 1, it is clear that the weighting term $\exp\{-j[\omega_S k + \nu_S \ell + \gamma_S]\}$ in (31) and (32) suppresses the oscillatory behavior of the sample moment. Since in our application we are interested in detecting the frequency of this oscillation and not in estimating the moments themselves, we shall use the statistic $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v^{(s+1)}(n,m)]$, which is expected to demonstrate an oscillatory behavior.

Since the principle of operation of the MPD operator $\overline{\mathrm{PD}}_{n^{(P)},m^{(s-P)}}$ and that of $\overline{\mathrm{PD}}_{n^{(0)},m^{(0)}}$ are identical (except that the first employs conjugated products, whereas the later uses unconjugated products for fields with a zero mean amplitude), the foregoing conclusions hold also for the problem of estimating $\overline{\mathrm{PD}}_{n^{(0)},m^{(0)}}[v(n,m)]$ from a single observed realization of a 2-D signal $v(n,m)$ of total-degree one. Hence, when the algorithm summarized in Table I is applied in practice, the $\overline{\mathrm{PD}}$ operators should be replaced by PD operators as concluded from (34).

V. Estimation in the Presence of Observation Noise

In Theorem 1, it is proved that in the absence of observation noise, the signal $\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[v(n,m)]$ is a 2-D exponential given by (18)–(20). Next, we show that a similar result holds for the more general case in which the observed signal consists of the sum of a random amplitude polynomial phase signal and additive white Gaussian noise (1)–(3).

Theorem 2: Let $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[y(n,m)]$ be the 2-D signal obtained by successively applying, in some arbitrary sequence, P times the operator $\mathrm{PD}_{n^{(1)}}[\cdot]$ and $S-P$ times the operator $\mathrm{PD}_{m^{(1)}}[\cdot]$ to the signal (1). Then, the signal $\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[y(n,m)]$ is the same 2-D exponential given by (18), i.e.,

$$\begin{aligned}\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[y(n,m)] &= \overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)] \\ &\quad \cdot \exp\{j[\omega_S n + \nu_S m + \gamma_S(\tau_n,\tau_m)]\} \\ n &= 0,1,\ldots,N-1-P\tau_n \\ m &= 0,1,\ldots,M-1-(S-P)\tau_m \end{aligned} \quad (36)$$

where ω_S and ν_S are given by (19) and (20), respectively. Both $\overline{\mathrm{PD}}_{n^{(P)},m^{(S-P)}}[w(n,m)]$ and $\gamma_S(\tau_n,\tau_m)$ are neither functions of m nor n.

Proof: Consider the 2-D signal $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[y(n,m)]$. From the recursive definition of the PD operator in Definition 1, as well as from (16), we note that for any complex valued field $\{y(n,m)\}$

$$\mathrm{PD}_{n^{(P)},m^{(S-P)}}[y(n,m)] = \prod_{q=0}^{S-P}\left\{\prod_{p=0}^{P}\left[y^{(*^{(p+q)})}(n+p\tau_n, m+q\tau_m)\right]^{\binom{P}{p}}\right\}^{\binom{S-P}{q}}. \quad (37)$$

Therefore, each shifted version $y(n+p\tau_n, m+q\tau_m)$ of the observed signal that appears in the product forms generated by applying the operator $\mathrm{PD}_{n^{(P)},m^{(S-P)}}[\cdot]$ to $y(n,m)$, is either conjugated or unconjugated but *never* appears in both forms for any S or P. Because $\{u(n,m)\}$ is a circular Gaussian white noise field, $E\{u^k(n,m)\}=0$ for any positive k, and $E\{u^k(n+\tau_n, m+\tau_m)[u^*(n,m)]^\ell\}=0$, unless $\tau_n=\tau_m=0$, and $k=\ell$. Moreover, since $\{u(n,m)\}$ is circular Gaussian and zero mean, all of its high-order moments can be expressed as functions of its second-order moments. Recall that $y(n,m)=v(n,m)+u(n,m)$. Because $\{w(n,m)\}$ and $\{u(n,m)\}$ are independent, and the second-order moments of shifted versions of $u(n,m)$ never involve both conjugated and unconjugated versions of the same sample, we conclude that all terms that involve high-order moments of $u(n,m)$ vanish. Hence

$$E\{\mathrm{PD}_{n^{(P)},m^{(S-P)}}[y(n,m)]\} = E\{\mathrm{PD}_{n^{(P)},m^{(S-P)}}[v(n,m)]\}. \quad (38)$$

□

We therefore conclude that the estimation algorithm derived in Section III can be applied *mutatis mutandis* to the estimation problem in the case of noisy observations. Moreover, following the arguments of Section IV, we conclude that by replacing ensemble averages with sample averages (i.e., replacing the $\overline{\mathrm{PD}}$ operators by PD operators), the same algorithm can be applied when only a single noisy realization of the observed field is available.

VI. CRB of a 2-D Random Amplitude Polynomial Phase Signal in Noise

In this section, we derive the CRB on the variance of the error in estimating the amplitude and phase parameters when the signal is observed in white additive Gaussian noise, i.e., the observed field is $\{y(n,m)\}$ given by (1)–(3). The amplitude $\{w(n,m)\}$ is assumed to be a real-valued, Gaussian field.

A. Problem Formulation

Define

$$\begin{aligned}\mathbf{y} = [&y(0,0),\ldots,y(0,M-1),y(1,0),\ldots \\ &y(1,M-1),\ldots,\ldots,y(N-1,0),\ldots \\ &y(N-1,M-1)]^T.\end{aligned} \quad (39)$$

The vector $\mathbf{w}$ is similarly defined. Let all the phase parameters of ϕ_S be assembled, layer by layer, into a vector $\mathbf{c}$

$$\mathbf{c} = [c(0,0); c(0,1), c(1,0); c(0,2), c(1,1) \\ c(2,0); \ldots, \ldots; c(0,S), \ldots, c(S,0)]^T \tag{40}$$

where we use ";" to distinguish layer from layer. Hence, $\mathbf{c}$ is an $\frac{(S+2)(S+1)}{2}$ dimensional vector.

In addition, let $\mathbf{t}$ be an $NM \times 2$ matrix such that each row of $\mathbf{t}$ contains a pair of indices (n,m), where $n = 0, \ldots, N-1$; $m = 0, \ldots, M-1$, and the rows are lexicographically ordered, i.e.,

$$\mathbf{t} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 & \ldots & 1 & \ldots & \ldots & N-1 \\ 0 & 1 & \ldots & M-1 & 0 & \ldots & M-1 & \ldots & \ldots & M-1 \end{bmatrix}^T. \tag{41}$$

We will use the following shorthand vector notation for functions of space: Given a scalar function $f(n,m)$, we denote the column vector consisting of the values of $f(n,m)$, $n = 0, \ldots, N-1$; $m = 0, \ldots, M-1$ by $f(\mathbf{t})$. Using this notation, we denote the vector of phase values of the signal by $\phi(\mathbf{t})$. Hence, we can define $\mathbf{A} = \text{diag}\{\cos\phi(\mathbf{t})\}$, $\mathbf{B} = \text{diag}\{\sin\phi(\mathbf{t})\}$, where $\cos\phi(\mathbf{t})$ and $\sin\phi(\mathbf{t})$ are MN-dimensional column vectors.

Let $\mathbf{y}^R = \text{Re}\{\mathbf{y}\}$, $\mathbf{y}^I = \text{Im}\{\mathbf{y}\}$, $\tilde{\mathbf{y}} = [\mathbf{y}^{R^T} \ \mathbf{y}^{I^T}]^T$. In a similar way, we define the noise vector $\tilde{\mathbf{u}}$. In this derivation, it is assumed that the amplitude field has a constant mean denoted by m_w. The covariance matrix of the vector $\mathbf{w}$ is denoted by $\mathbf{R}_w$ and is assumed to have some known parametric form, where $\mathbf{a}$ is the parameter vector. At the moment, we will not specify the functional dependence of $\mathbf{R}_w$ on $\mathbf{a}$, but rather leave it implicit. As an example, we may assume that it is the covariance matrix of a finite-dimensional moving-average (MA) field parameterized by the MA model coefficients.

The observation noise $u(n,m)$ is an additive complex valued, zero mean, circular white Gaussian noise of unknown variance σ^2. Hence, the noise field can be written as $u(n,m) = u_1(n,m) + ju_2(n,m)$, with $u_1(n,m)$, and $u_2(n,m)$ being independent, identically distributed, real-valued white Gaussian noise fields, with variance $\sigma^2/2$ each. Both $u_1(n,m)$ and $u_2(n,m)$ are assumed to be independent of the amplitude function $w(n,m)$.

Finally, we collect all of the unknown parameters into a single vector $\boldsymbol{\theta}$, such that

$$\boldsymbol{\theta} = \{\mathbf{c}, m_w, \mathbf{a}, \sigma^2\}. \tag{42}$$

The problem considered in this section can now be stated as follows. Given the measurements $\{y(n,m)\}$, how accurately can the parameter vector $\boldsymbol{\theta}$ be estimated?

B. Derivation of the CRB

Rewriting the measurements equation (1) in a vector form using real quantities only, we have

$$\tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathbf{w} + \tilde{\mathbf{u}}. \tag{43}$$

Since $\mathbf{u}^R$, $\mathbf{u}^I$, and $\mathbf{w}$ are Gaussian and independent, $\tilde{\mathbf{y}}$ is Gaussian as well. Let $\boldsymbol{\mu}$ denote the mean vector of $\tilde{\mathbf{y}}$, let $\boldsymbol{\Gamma}$ denote its covariance matrix, and let

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}. \tag{44}$$

Thus

$$\boldsymbol{\mu} = m_w \mathbf{X} \mathbf{1}_{NM} \tag{45}$$

and

$$\boldsymbol{\Gamma} = \mathbf{X}\mathbf{R}_w\mathbf{X}^T + \frac{\sigma^2}{2}\mathbf{I}_{2NM} \tag{46}$$

where $\mathbf{1}_{NM}$ is an NM-dimensional column vector of ones, and $\mathbf{I}_{2NM}$ is a $2NM$ identity matrix.

Let

$$\mathbf{e}_1 = [0, 1, \ldots, (N-1)]^T \otimes \mathbf{1}_M \tag{47}$$

$$\mathbf{e}_2 = \mathbf{1}_N \otimes [0, 1, \ldots, (M-1)]^T \tag{48}$$

where $\mathbf{1}_M$ and $\mathbf{1}_N$ are M-dimensional and N-dimensional column vectors of ones, respectively, and $\otimes$ is the Kronecker product. In other words, $\mathbf{e}_1$ is the first column of $\mathbf{t}$, and $\mathbf{e}_2$ is its second column. Define

$$\mathbf{T}_N^k = (\text{diag}\{\mathbf{e}_1\})^k \tag{49}$$

and similarly

$$\mathbf{T}_M^\ell = (\text{diag}\{\mathbf{e}_2\})^\ell. \tag{50}$$

Let Λ denote the log-likelihood function of the observation vector $\tilde{\mathbf{y}}$. The general expression for the Fisher information matrix (FIM) of a real valued Gaussian process is given by (e.g., [15])

$$-E\left\{\frac{\partial^2 \Lambda}{\partial\boldsymbol{\theta}(k)\partial\boldsymbol{\theta}(\ell)}\right\} = \frac{\partial\boldsymbol{\mu}^T}{\partial\boldsymbol{\theta}(k)}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}(\ell)} + \frac{1}{2}\text{tr}\left\{\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\boldsymbol{\theta}(k)}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\boldsymbol{\theta}(\ell)}\right\}. \tag{51}$$

To evaluate (51), we need to compute the derivatives of $\boldsymbol{\mu}$ and $\boldsymbol{\Gamma}$ with respect to the various parameters of interest:

$$\frac{\partial\boldsymbol{\mu}}{\partial m_w} = \mathbf{X}\mathbf{1}_{NM} \tag{52}$$

$$\frac{\partial\boldsymbol{\mu}}{\partial c(k,\ell)} = m_w\frac{\partial\mathbf{X}}{\partial c(k,\ell)}\mathbf{1}_{NM} \tag{53}$$

$$\frac{\partial\boldsymbol{\mu}}{\partial a(k)} = 0 \tag{54}$$

$$\frac{\partial\boldsymbol{\mu}}{\partial\sigma^2} = 0 \tag{55}$$

$$\frac{\partial\boldsymbol{\Gamma}}{\partial m_w} = 0 \tag{56}$$

$$\frac{\partial\boldsymbol{\Gamma}}{\partial c(k,\ell)} = \frac{\partial\mathbf{X}}{\partial c(k,\ell)}\mathbf{R}_w\mathbf{X}^T + \mathbf{X}\mathbf{R}_w\frac{\partial\mathbf{X}^T}{\partial c(k,\ell)} \tag{57}$$

$$\frac{\partial\boldsymbol{\Gamma}}{\partial a(k)} = \mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T \tag{58}$$

$$\frac{\partial\boldsymbol{\Gamma}}{\partial\sigma^2} = \frac{1}{2}\mathbf{I}_{2NM}. \tag{59}$$

We thus immediately conclude that

$$-E\left\{\frac{\partial^2\Lambda}{\partial a(k)\partial m_w}\right\}=0 \tag{60}$$

and

$$-E\left\{\frac{\partial^2\Lambda}{\partial\sigma^2\partial m_w}\right\}=0. \tag{61}$$

Taking the partial derivatives with respect to the phase parameters yields

$$\frac{\partial\mathbf{X}}{\partial c(k,\ell)}=\mathbf{H}_{k,\ell}\mathbf{V} \tag{62}$$

where we define

$$\mathbf{H}_{k,\ell}=\begin{bmatrix}\mathbf{T}_N^k\mathbf{T}_M^\ell & 0\\ 0 & \mathbf{T}_N^k\mathbf{T}_M^\ell\end{bmatrix} \tag{63}$$

$$\mathbf{V}=\begin{bmatrix}-\mathbf{B}\\ \mathbf{A}\end{bmatrix}. \tag{64}$$

Substituting (62) into (53) and (57), we have

$$\frac{\partial\boldsymbol{\mu}}{\partial c(k,\ell)}=m_w\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{1}_{NM} \tag{65}$$

$$\frac{\partial\boldsymbol{\Gamma}}{\partial c(k,\ell)}=\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T+\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}. \tag{66}$$

Next, we will use the following identities:

$$\mathbf{X}^T\mathbf{X}=\mathbf{V}^T\mathbf{V}=\mathbf{I}_{NM} \tag{67}$$

$$\mathbf{X}^T\mathbf{V}=\mathbf{V}^T\mathbf{X}=0 \tag{68}$$

$$\mathbf{X}\mathbf{X}^T+\mathbf{V}\mathbf{V}^T=\mathbf{I}_{2NM} \tag{69}$$

$$\mathbf{X}^T\mathbf{H}_{k,\ell}\mathbf{V}=\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}=0 \tag{70}$$

$$\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{V}=\mathbf{T}_N^k\mathbf{T}_M^\ell \tag{71}$$

$$\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}=\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}. \tag{72}$$

Using (46) and the matrix inversion lemma (e.g., [8]), we find that

$$\begin{aligned}\boldsymbol{\Gamma}^{-1}&=\frac{2}{\sigma^2}\mathbf{I}_{2NM}-\frac{2}{\sigma^2}\mathbf{X}\left(\frac{2}{\sigma^2}\mathbf{X}^T\mathbf{X}+\mathbf{R}_w^{-1}\right)^{-1}\mathbf{X}^T\frac{2}{\sigma^2}\\ &=\frac{2}{\sigma^2}\mathbf{I}_{2NM}-\frac{2}{\sigma^2}\mathbf{X}\left(\mathbf{I}_{NM}+\frac{\sigma^2}{2}\mathbf{R}_w^{-1}\right)^{-1}\mathbf{X}^T\\ &=\frac{2}{\sigma^2}\mathbf{I}_{2NM}-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\end{aligned} \tag{73}$$

where the second equality results from (67), and we define

$$\mathbf{D}=\mathbf{I}_{NM}+\frac{\sigma^2}{2}\mathbf{R}_w^{-1}. \tag{74}$$

Using (73) and (66), we have that

$$\begin{aligned}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial c(k,\ell)}&=\left(\frac{2}{\sigma^2}\mathbf{I}_{2NM}-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\right)\\ &\quad\times(\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T+\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell})\\ &=\frac{2}{\sigma^2}\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T+\frac{2}{\sigma^2}\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\\ &\quad-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\\ &\quad-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\\ &=\frac{2}{\sigma^2}\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T+\frac{2}{\sigma^2}\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\\ &\quad-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\end{aligned} \tag{75}$$

where the third equality results from (67) and (70). Similarly, using (73) and (58), we have that

$$\begin{aligned}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial a(k)}&=\left(\frac{2}{\sigma^2}\mathbf{I}_{2NM}-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\right)\left(\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\right)\\ &=\frac{2}{\sigma^2}\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T-\frac{2}{\sigma^2}\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T.\end{aligned} \tag{76}$$

Using (51) and (54), we have

$$\begin{aligned}&-E\left\{\frac{\partial^2\Lambda}{\partial c(k,\ell)\partial a(p)}\right\}\\ &=\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\left\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(p)}\mathbf{X}^T\right\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\left\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(p)}\mathbf{X}^T\right\}\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\left\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(p)}\mathbf{X}^T\right\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\left\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(p)}\mathbf{X}^T\right\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\left\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(p)}\mathbf{X}^T\right\}\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\left\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(p)}\mathbf{X}^T\right\}\\ &=0\end{aligned} \tag{77}$$

where we have used (70) and the commutative property of the trace operator. Using similar arguments, we also have

$$\begin{aligned}&-E\left\{\frac{\partial^2\Lambda}{\partial c(k,\ell)\partial c(p,q)}\right\}\\ &=m_w^2\left[\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}\boldsymbol{\Gamma}^{-1}\mathbf{H}_{p,q}\mathbf{V}\mathbf{1}_{NM}\right]\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\mathbf{H}_{p,q}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\}\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\}\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\}\\ &\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\ &\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\mathrm{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\end{aligned}$$

$$
\begin{aligned}
&= \frac{2m_w^2}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}(\mathbf{I}_{2NM}-\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T)\mathbf{H}_{p,q}\mathbf{V}\mathbf{1}_{NM}\right]\\
&\quad+\frac{2}{\sigma^4}\operatorname{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\
&\quad-\frac{2}{\sigma^4}\operatorname{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\
&\quad+\frac{2}{\sigma^4}\operatorname{tr}\{\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}\mathbf{R}_w\}\\
&\quad-\frac{2}{\sigma^4}\operatorname{tr}\{\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}\mathbf{R}_w\}\\
&= \frac{2m_w^2}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}\mathbf{1}_{NM}\right.\\
&\quad\left.-\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\mathbf{H}_{p,q}\mathbf{V}\mathbf{1}_{NM}\right]\\
&\quad+\frac{2}{\sigma^4}\operatorname{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w(I_{NM}-\mathbf{D}^{-1})\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\
&\quad+\frac{2}{\sigma^4}\operatorname{tr}\{\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{H}_{p,q}\mathbf{V}\mathbf{R}_w(I_{NM}-\mathbf{D}^{-1})\}\\
&= \frac{2m_w^2}{\sigma^2}\mathbf{1}_{NM}^T\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}\mathbf{1}_{NM}\\
&\quad+\frac{4}{\sigma^4}\operatorname{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w(I_{NM}-\mathbf{D}^{-1})\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{p,q}\}\\
&= \frac{2m_w^2}{\sigma^2}\sum_{n=0}^{N-1}n^{k+p}\sum_{m=0}^{M-1}m^{\ell+q}\\
&\quad+\frac{4}{\sigma^4}\operatorname{tr}\{\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}\mathbf{R}_w(I_{NM}-\mathbf{D}^{-1})\mathbf{R}_w\}
\end{aligned}
\tag{78}
$$

where we have used (67) and (70); the commutative property of the trace operator; the symmetric property of $\mathbf{H}_{k,\ell}$, $\mathbf{H}_{p,q}$, and $\mathbf{R}_w$; and the diagonality of $\mathbf{H}_{k,\ell}$ and $\mathbf{H}_{p,q}$. The last equality is due to (72).

Using (52), (56), (65), and (70), we have

$$
\begin{aligned}
-E&\left\{\frac{\partial^2\Lambda}{\partial c(k,\ell)\partial m_w}\right\}\\
&= \frac{\partial\boldsymbol{\mu}^T}{\partial c(k,\ell)}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\mu}}{\partial m_w}\\
&= \frac{2m_w}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}(\mathbf{I}_{2NM}-\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T)\mathbf{X}\mathbf{1}_{NM}\right]\\
&= \frac{2m_w}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{1}_{NM}\right.\\
&\quad\left.-\mathbf{1}_{NM}^T\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\mathbf{X}\mathbf{1}_{NM}\right]\\
&= 0.
\end{aligned}
\tag{79}
$$

Additionally, using (52), (56), and (67), we find that

$$
\begin{aligned}
-E&\left\{\frac{\partial^2\Lambda}{\partial^2 m_w}\right\}\\
&= \frac{\partial\boldsymbol{\mu}}{\partial m_w}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\mu}^T}{\partial m_w}\\
&= \frac{2}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{X}^T(\mathbf{I}_{2NM}-\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T)\mathbf{X}\mathbf{1}_{NM}\right]\\
&= \frac{2}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{X}^T\mathbf{X}\mathbf{1}_{NM}-\mathbf{1}_{NM}^T\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\mathbf{X}\mathbf{1}_{NM}\right]\\
&= \frac{2}{\sigma^2}\left[\mathbf{1}_{NM}^T\mathbf{1}_{NM}-\mathbf{1}_{NM}^T\mathbf{D}^{-1}\mathbf{1}_{NM}\right]\\
&= \frac{2}{\sigma^2}\left[NM-\sum_{k=1}^{NM}\sum_{\ell=1}^{NM}\mathbf{D}^{-1}(k,\ell)\right].
\end{aligned}
\tag{80}
$$

Using (54), (67), the trace operator commutative property, the invariance of the trace operator to transposition, and the symmetric property of $\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}$ and $\mathbf{D}^{-1}$, we obtain

$$
\begin{aligned}
-E&\left\{\frac{\partial^2\Lambda}{\partial a(k)\partial a(\ell)}\right\}\\
&= \frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\mathbf{X}^T\right\}\\
&\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\mathbf{X}^T\right\}\\
&\quad-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\mathbf{X}^T\right\}\\
&\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\mathbf{X}^T\right\}\\
&= \frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\right\}\\
&\quad-\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\right\}\\
&\quad+\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(\ell)}\right\}.
\end{aligned}
\tag{81}
$$

Finally, substituting (55), (59) and (75) into (51) while using (70), and the commutativity of the trace operator, gives

$$
\begin{aligned}
-E&\left\{\frac{\partial^2\Lambda}{\partial c(k,\ell)\partial\sigma^2}\right\}\\
&= \left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\}\\
&\quad-\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\{\mathbf{H}_{k,\ell}\mathbf{V}\mathbf{R}_w\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\}\\
&\quad+\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\}\\
&\quad-\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\{\mathbf{X}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\}\\
&\quad-\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\}\\
&\quad+\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{R}_w\mathbf{V}^T\mathbf{H}_{k,\ell}\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\}\\
&= 0.
\end{aligned}
\tag{82}
$$

In addition

$$
\begin{aligned}
-E&\left\{\frac{\partial^2\Lambda}{\partial a(k)\partial\sigma^2}\right\}\\
&= \left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\right\}\\
&\quad-\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\right\}\\
&\quad-\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\right\}\\
&\quad+\left(\frac{1}{\sigma^2}\right)^2\operatorname{tr}\left\{\mathbf{X}\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\right\}\\
&= \frac{1}{\sigma^4}\operatorname{tr}\left\{\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\right\}-\frac{2}{\sigma^4}\operatorname{tr}\left\{\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{D}^{-1}\right\}\\
&\quad+\frac{1}{\sigma^4}\operatorname{tr}\left\{\mathbf{D}^{-1}\frac{\partial\mathbf{R}_w}{\partial\mathbf{a}(k)}\mathbf{D}^{-1}\right\}.
\end{aligned}
\tag{83}
$$

The FIM entry that corresponds to the noise parameter is given by

$$-E\left\{\frac{\partial^2\Lambda}{\partial\sigma^2\partial\sigma^2}\right\} = \frac{1}{2}\left(\frac{1}{\sigma^2}\right)^2 \mathrm{tr}\{\mathbf{I}_{2NM}\} - \left(\frac{1}{\sigma^2}\right)^2 \mathrm{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\} + \frac{1}{2}\left(\frac{1}{\sigma^2}\right)^2 \mathrm{tr}\{\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\mathbf{X}\mathbf{D}^{-1}\mathbf{X}^T\} = \frac{NM}{\sigma^4} - \frac{1}{\sigma^4}\mathrm{tr}\{\mathbf{D}^{-1}\} + \frac{1}{2\sigma^4}\mathrm{tr}\{\mathbf{D}^{-1}\mathbf{D}^{-1}\}. \tag{84}$$

We can now summarize our observations regarding the CRB for a 2-D random amplitude polynomial phase signal. The bounds on the parameter estimates of the phase, the amplitude mean, and the parameter vector $\mathbf{a}$ of the amplitude covariance function are mutually decoupled. Moreover, the elements of the FIM are independent of the specific model of the amplitude field since those of them which depend on the amplitude are functions of its mean and covariance matrix $\mathbf{R}_w$ only. Thus, closed-form formulas for the CRB are obtained by substituting into $\mathbf{R}_w$ the expression for the amplitude field covariance matrix, which is expressed in terms of the amplitude field parameters. As an example, the expression for the covariance matrix of a nonsymmetric half-plane 2-D moving average field is derived in [16].

Because the CRB for the phase is decoupled from the bound on the amplitude and noise parameters, it can be obtained by inverting (78). The CRB for the phase parameters is independent of the specific parametric model used for the covariance of the amplitude field, as well as of the specific values of the phase parameters. Thus, all signals whose phase is of some total degree S, and whose amplitude have the same mean and covariance functions, will have identical values for the CRB on the phase parameters. The bounds on the amplitude parameters and the noise variance are both independent and decoupled from the phase.

The bound on the mean of the amplitude field is decoupled from the bounds on the other parameters. It is a function only of the amplitude covariance and the observation noise variance. The CRB on the parameters of the amplitude covariance is independent of the phase function and of the amplitude mean but is a function of the observation noise variance and the amplitude covariance. Note that the FIM block corresponding to the parameters of the amplitude covariance and the observation noise is *identical* to the block we would have obtained if the amplitude field was not modulated by $e^{j\phi(n,m)}$. Hence, the CRB for these amplitude parameters is the same as if the modulation by $e^{j\phi(n,m)}$ was not present. Similarly, the bound on the noise variance is also decoupled from the bounds on the phase and mean and is identical to the bound obtained when the modulation by $e^{j\phi(n,m)}$ is not present.

Finally, we note that in many cases, we are interested not in the phase or amplitude parameters themselves but in estimating some function of these parameters. For example, having estimated the model parameters, the amplitude field spectral density and the local phase and frequency functions can be computed using their known functional dependence on the (estimated) parameters. Next, we derive the CRB on the local phase and frequency functions.

Since the local phase defined in (3) is a differentiable function of the phase parameter vector $\mathbf{c}$, the CRB on $\phi_S(n,m)$ is related to the CRB of $\mathbf{c}$ by (e.g., [17])

$$\mathrm{CRB}(\phi_S(n,m)) = \mathbf{g}^T\mathrm{CRB}(\mathbf{c})\mathbf{g} \tag{85}$$

where

$$\mathbf{g} = \left[\frac{\partial\phi_S(n,m)}{\partial c(0,0)}, \frac{\partial\phi_S(n,m)}{\partial c(0,1)}, \frac{\partial\phi_S(n,m)}{\partial c(1,0)}, \ldots \frac{\partial\phi_S(n,m)}{\partial c(S-1,1)} \frac{\partial\phi_S(n,m)}{\partial c(S,0)}\right]^T = [1, m, n, \ldots, n^{S-1}m, n^S]^T. \tag{86}$$

In the case of continuous index fields, the local spatial frequencies are the partial derivatives of the local phase function. Thus, assuming for a moment n and m to be continuous variables, we have

$$\omega(n,m) = \frac{1}{2\pi}\frac{\phi_S(n,m)}{\partial n} = \frac{1}{2\pi}\sum_{(k,\ell)\in\{1\le k;0\le\ell;1\le k+\ell\le S\}} c(k,\ell)kn^{k-1}m^{\ell} \tag{87}$$

and

$$\nu(n,m) = \frac{1}{2\pi}\frac{\phi_S(n,m)}{\partial m} = \frac{1}{2\pi}\sum_{(k,\ell)\in\{0\le k;1\le\ell;1\le k+\ell\le S\}} c(k,\ell)\ell n^{k}m^{\ell-1}. \tag{88}$$

Hence

$$\mathrm{CRB}(\omega(n,m)) = \mathbf{h}_n^T\mathrm{CRB}(\mathbf{c})\mathbf{h}_n \tag{89}$$

where

$$\mathbf{h}_n = \left[\frac{\partial\omega(n,m)}{\partial c(0,0)}, \frac{\partial\omega(n,m)}{\partial c(0,1)}, \frac{\partial\omega(n,m)}{\partial c(1,0)}, \ldots \frac{\partial\omega(n,m)}{\partial c(S-1,1)}, \frac{\partial\omega(n,m)}{\partial c(S,0)}\right]^T \tag{90}$$

and

$$\frac{\partial\omega(n,m)}{\partial c(k,\ell)} = \frac{1}{2\pi}kn^{k-1}m^{\ell} \qquad 1\le k;\quad 0\le \ell;\quad 1\le k+\ell\le S. \tag{91}$$

Similarly

$$\mathrm{CRB}(\nu(n,m)) = \mathbf{h}_m^T\mathrm{CRB}(\mathbf{c})\mathbf{h}_m \tag{92}$$

where

$$\mathbf{h}_m = \left[\frac{\partial\nu(n,m)}{\partial c(0,0)}, \frac{\partial\nu(n,m)}{\partial c(0,1)}, \frac{\partial\nu(n,m)}{\partial c(1,0)}, \ldots \frac{\partial\nu(n,m)}{\partial c(S-1,1)} \frac{\partial\nu(n,m)}{\partial c(S,0)}\right]^T \tag{93}$$

and

$$\frac{\partial \nu(n,m)}{\partial c(k,\ell)} = \frac{1}{2\pi}\ell n^k m^{\ell-1}$$
$$0 \le k; \quad 1 \le \ell; \quad 1 \le k+\ell \le S. \quad (94)$$

C. CRB for High SNR

In this section, we specialize the general results derived in the previous section for the case where the measurements of the signal are known have high SNR. In other words, we assume here that $\sigma^2 \to 0$. Hence, a first-order approximation of $\mathbf{D}^{-1}$ yields

$$\mathbf{D}^{-1} \approx \mathbf{I}_{NM} - \frac{\sigma^2}{2}\mathbf{R}_w^{-1}. \quad (95)$$

Thus, (74) can be approximated by

$$\begin{aligned}\mathbf{\Gamma}^{-1} &\approx \frac{2}{\sigma^2}\mathbf{I}_{2NM} - \frac{2}{\sigma^2}\mathbf{X}\left(\mathbf{I}_{NM} - \frac{\sigma^2}{2}\mathbf{R}_w^{-1}\right)\mathbf{X}^T \\ &= \frac{2}{\sigma^2}(\mathbf{I}_{2NM} - \mathbf{X}\mathbf{X}^T) + \mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T \\ &= \frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T + \mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T. \quad (96)\end{aligned}$$

Substituting (95) and (96) into the equations of the nonzero elements of the FIM yields the FIM for the high SNR case. In particular, substituting (95) into (78), we have

$$\begin{aligned}&-E\left\{\frac{\partial^2 \Lambda}{\partial c(k,\ell)\partial c(p,q)}\right\} \\ &= \frac{2m_w^2}{\sigma^2}\sum_{n=0}^{N-1} n^{k+p}\sum_{m=0}^{M-1} m^{\ell+q} \\ &\quad + \frac{4}{\sigma^4}\,\mathrm{tr}\left\{\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}\mathbf{R}_w\frac{\sigma^2}{2}\mathbf{R}_w^{-1}\mathbf{R}_w\right\} \\ &= \frac{2m_w^2}{\sigma^2}\sum_{n=0}^{N-1} n^{k+p}\sum_{m=0}^{M-1} m^{\ell+q} + \frac{2}{\sigma^2}\,\mathrm{tr}\{\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}\mathbf{R}_w\} \\ &= \frac{2m_w^2}{\sigma^2}\sum_{n=0}^{N-1} n^{k+p}\sum_{m=0}^{M-1} m^{\ell+q} + \frac{2r_w(0,0)}{\sigma^2}\,\mathrm{tr}\{\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}\} \\ &= \frac{2(m_w^2 + r_w(0,0))}{\sigma^2}\sum_{n=0}^{N-1} n^{k+p}\sum_{m=0}^{M-1} m^{\ell+q} \quad (97)\end{aligned}$$

where the third equality is due to the diagonality of $\mathbf{T}_N^{k+p}\mathbf{T}_M^{\ell+q}$ and since all the elements of the main diagonal of $\mathbf{R}_w$ are equal to $r_w(0,0)$. Here, $r_w(0,0)$ denotes the variance of the amplitude field. Let $\mathrm{SNR} = \frac{m_w^2 + r_w(0,0)}{\sigma^2}$ denote the signal-to-noise ratio. Using (97), we conclude that for high SNR scenarios, the CRB on the error variance in estimating the phase parameters is inversely proportional to the SNR. Substituting (95) into (80), we get

$$-E\left\{\frac{\partial^2 \Lambda}{\partial^2 m_w}\right\} = \sum_{k=1}^{NM}\sum_{\ell=1}^{NM}\mathbf{R}_w^{-1}(k,\ell). \quad (98)$$

Using (58) and (96), we have

$$\begin{aligned}\mathbf{\Gamma}^{-1}\frac{\partial \mathbf{\Gamma}}{\partial a(k)} &= \left(\frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T + \mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\right)\left(\mathbf{X}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T\right) \\ &= \frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T\mathbf{X}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T + \mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\mathbf{X}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T \\ &= \mathbf{X}\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T \quad (99)\end{aligned}$$

where the last equality follows from (67) and (68). Hence

$$\begin{aligned}&-E\left\{\frac{\partial^2 \Lambda}{\partial a(k)\partial a(\ell)}\right\} \\ &= \frac{1}{2}\,\mathrm{tr}\left\{\mathbf{X}\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(\ell)}\mathbf{X}^T\right\} \\ &= \frac{1}{2}\,\mathrm{tr}\left\{\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(\ell)}\right\}. \quad (100)\end{aligned}$$

Note that (100) is *identical* to the expression we would have obtained if the amplitude field was zero-mean and was measured directly (i.e., if the modulation by $e^{j\phi(n,m)}$ did not take place and the observations were noise free). Thus, we can use here any available expression for the FIM of a real valued, zero mean, homogeneous Gaussian random field. For example, if the amplitude was a nonsymmetric half-plane (NSHP) moving average field, we could use the expressions derived in [16].

The FIM entry that corresponds to the noise parameter is given by

$$\begin{aligned}-E\left\{\frac{\partial^2 \Lambda}{\partial \sigma^2 \partial \sigma^2}\right\} &= \frac{1}{8}\,\mathrm{tr}\left\{\frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T\frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T\right\} \\ &\quad + \frac{1}{8}\,\mathrm{tr}\left\{\frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T\mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\right\} \\ &\quad + \frac{1}{8}\,\mathrm{tr}\left\{\mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T\right\} \\ &\quad + \frac{1}{8}\,\mathrm{tr}\{\mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\} \\ &= \frac{1}{2\sigma^4}\,\mathrm{tr}\{(\mathbf{V}\mathbf{V}^T)^2\} + \frac{1}{8}\,\mathrm{tr}\{\mathbf{R}_w^{-1}\mathbf{R}_w^{-1}\} \\ &= \frac{NM}{2\sigma^4} + \frac{1}{8}\,\mathrm{tr}\{\mathbf{R}_w^{-1}\mathbf{R}_w^{-1}\}. \quad (101)\end{aligned}$$

Similarly

$$\begin{aligned}-E\left\{\frac{\partial^2 \Lambda}{\partial a(k)\partial \sigma^2}\right\} &= \frac{1}{4}\,\mathrm{tr}\left\{\mathbf{X}\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T\frac{2}{\sigma^2}\mathbf{V}\mathbf{V}^T\right\} \\ &\quad + \frac{1}{4}\,\mathrm{tr}\left\{\mathbf{X}\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{X}^T\mathbf{X}\mathbf{R}_w^{-1}\mathbf{X}^T\right\} \\ &= \frac{1}{4}\,\mathrm{tr}\left\{\mathbf{R}_w^{-1}\frac{\partial \mathbf{R}_w}{\partial \mathbf{a}(k)}\mathbf{R}_w^{-1}\right\} \\ &= -\frac{1}{4}\,\mathrm{tr}\left\{\frac{\partial \mathbf{R}_w^{-1}}{\partial \mathbf{a}(k)}\right\}. \quad (102)\end{aligned}$$

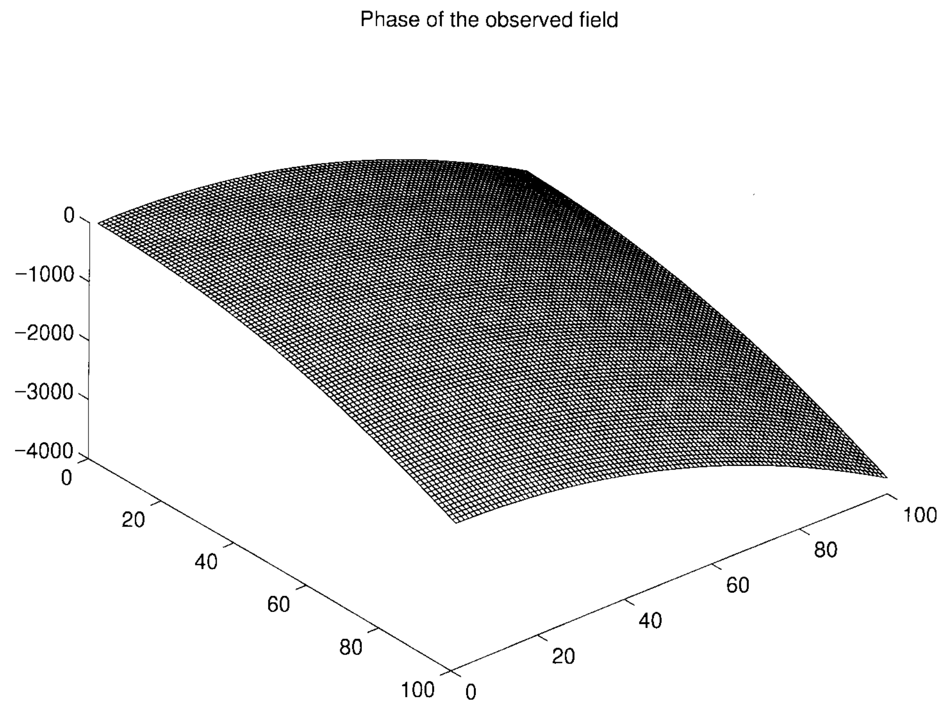


Fig. 2. True phase function of the observed signal.

Note that as σ^2 tends to zero, the FIM block (97), which corresponds to the phase parameters, becomes singular, and hence, the phase of the signal can be perfectly estimated, regardless of the structure of amplitude covariance matrix. This result is due to the fact that in the absence of observation noise, the phase of the measured signal $y(n,m)$ can be obtained by dividing the imaginary part of the measured signal by its real part.

VII. Numerical Examples

To illustrate the operation of the proposed algorithm, as well as to gain more insight into its performance relative to the CRB, we present numerical evaluation for some specific examples.

Example 1: Consider a random amplitude polynomial phase signal of total-degree 2. The amplitude is exponentially distributed with parameter $\lambda = 1$ [i.e., the amplitude field samples are i.i.d. random variables with probability density function given by $p(w(n,m)) = \frac{1}{\lambda}\exp\{-w(n,m)/\lambda\}$]. The observations are subject to an additive complex valued, white Gaussian noise, such that the SNR $= -3$ dB. In this case, the SNR is defined as SNR $= 10\log\frac{\lambda^2}{\sigma^2}$, where σ^2 is the variance of the additive noise. In this example, the observed field dimensions are $N = 100$ and $M = 100$. The phase coefficients are given by $\mathbf{c} = [-1.5; 0.311, 0.211; -0.1555, -0.01, -0.22]^T$.

The true phase function is shown in Fig. 2. Note the very low sampling rate of this phase function (the phase-axis of this figure is measured in radians, and the dimensions of the sampling grid are 100×100). Many of the existing phase estimation and restoration algorithms are adversely affected by insufficient spatial sampling (with respect to the instantaneous frequency) and noise; see, e.g., [12].

Since the amplitude field is positive, we employ the estimation algorithm derived in Section III-A. Next, we illustrate the steps of the estimation algorithm. Since the polynomial phase total-degree is 2, we start by estimating the parameters of layer 2. In the first step of the algorithm, we have $s = 1$ and $P = 0$. Hence, applying the operator $\text{PD}_{n^{(0)},m^{(1)}}$ to the observed signal, we obtain the signal denoted by $x(n,m)$, which is (approximately, due to the noise) a 2-D random-amplitude polynomial phase signal of total-degree 1, i.e., a 2-D random-amplitude exponential. The absolute value of this signal (which is the amplitude field, except for the contributions of the observation noise) is shown in the left image of Fig. 3. The absolute value of the signal DFT is shown in the right-hand side of the same figure. Note that although the observed field is undersampled and the noise level is high, applying the proposed operator to the observed signal results in a prominent spectral peak. Estimating the spatial frequency of the spectral peak results in the estimates of $c(1,1)$ and $c(0,2)$.

Repeating the same procedure for $s = 1$ and $P = 1$, i.e., applying the operator $\text{PD}_{n^{(1)},m^{(0)}}$ to the observed signal, we obtain another 2-D random amplitude exponential signal. Estimating the spatial frequency of the spectral peak results in the estimates of $c(2,0)$ and $c(1,1)$. We have therefore obtained estimates for all three parameters of layer 2. Multiplying $y(n,m)$ by $\exp\{-j\psi_2(n,m)\}$, we obtain a new, approximately polynomial-phase signal with random amplitude $y^{(1)}(n,m)$, whose total degree is 1. Since in this iteration $s = 0$ and $P = 0$, the parameters $c(1,0)$ and $c(0,1)$ of layer 1 are estimated by finding the spatial frequency of the peak of the 2-D signal DFT. Multiplying $y^{(1)}(n,m)$ by $\exp\{-j\psi_1(n,m)\}$, we obtain the signal $y^{(0)}(n,m)$, whose total degree is 0. The coefficient $c(0,0)$ can now be computed as the arithmetic average of the imaginary part of the logarithm of $y^{(0)}(n,m)$.

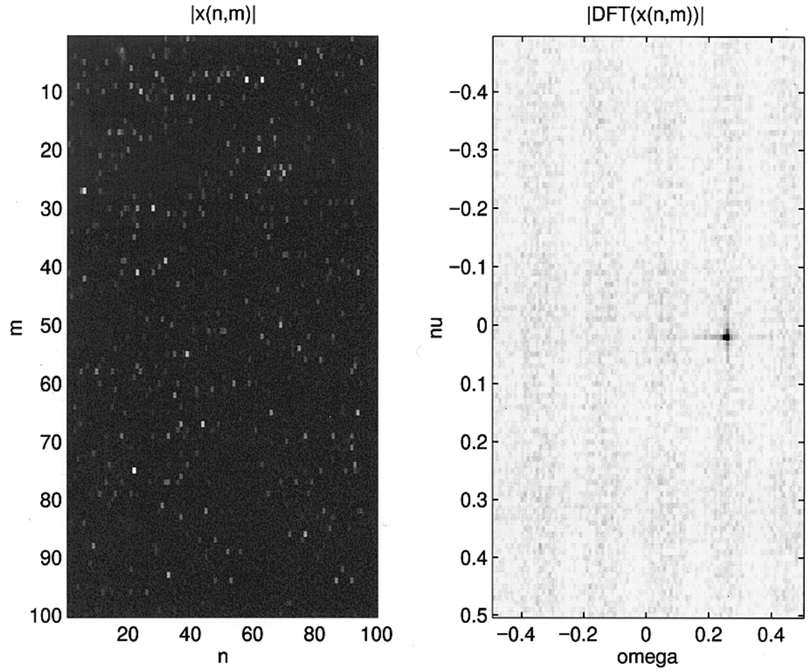


Fig. 3. Two-dimensional random amplitudepolynomial phase signal after applying the operator $\mathrm{PD}_{n^{(0)},m^{(1)}}$ to the observed signal of total-degree 2. (In this iteration, $s = 1$, $p = 0$, and the signal is the observed signal $y^{(2)}(n, m)$). Left: Absolute value of the resulting 2-D signal. Right: Absolute value of the resulting 2-D signal DFT.

Thus, we have completed the estimation of all the phase parameters of the observed 2-D nonhomogeneous signal.

Example 2: In this example, we illustrate the performance of the algorithm by Monte Carlo simulations. We analyze the bias and the variance of the estimates obtained by the algorithm and compare them with the CRB, which was derived in Section VI-C, under the high SNR assumption.

In these examples, the observation noise is a complex valued, zero mean, white Gaussian noise, and we investigate the performance of the algorithm as a function of the SNR. The random amplitude of the polynomial phase signal is a 2-D Gaussian, NSHP moving average field, with mean equal to 10. The NSHP moving average model has an $S_{3,3}$ support.

The Fourier transform of the covariance function of the MA field is depicted in Fig. 4. For this field, the ratio $r_w(1,0)/r_w(0,0) = 0.4469$. The phase function of the 2-D signal is of total-degree 2. The phase parameter vector is given by $\mathbf{c} = [2; 0.045, 0.082; -0.0015, 0.0016, -0.0022]^T$. The observed field dimensions are 100×100. The experimental results are based on 300 independent realizations of the observed signal for each SNR value. Note that here, the SNR is defined as $\mathrm{SNR} = 10 \log \frac{m_w^2 + r_w(0,0)}{\sigma^2}$, where $r_w(0,0)$ is the MA field variance, and σ^2 is the variance of the additive noise. In this example, the SNR varies by changing the observation noise variance from experiment to experiment, whereas m_w and $r_w(0,0)$ are held fixed.

In order to demonstrate the crucial importance of the choice of the algorithm parameters τ_n and τ_m, we have repeated the Monte Carlo experiments for two different sets of these parameters. In the first case, τ_n and τ_m were chosen to be relatively large ($\tau_n = \frac{N}{2}, \tau_m = \frac{M}{2}$). These parameters were set to small values in the second experiment, where we chose $\tau_n = \tau_m = 1$. The Monte Carlo simulations indicate that the estimator proposed in Section III-A yields unbiased estimates of the phase parameters, as the experimental bias is considerably smaller than the experimental standard deviation [except for the estimates of $c(0,0)$, which are slightly biased]. The estimation error variance can therefore be compared with the CRB derived in Section VI. (The CRB provides the lower bound on the error variance for any unbiased estimator of the problem parameters). A comparison of the Monte Carlo results with the CRB, which is computed assuming a high SNR, is depicted in Fig. 5. The experimental results for the selection of $\tau_n = \frac{N}{2}$ and $\tau_m = \frac{M}{2}$ (dashed-dotted line), indicate that the phase estimates are 7–10 dB above the high SNR CRB. Note that this result holds for low SNR values as well, although the high SNR assumption used to compute the bound is not valid anymore, and the bound should be considered to be an optimistic one. However, for the selection of $\tau_n = \tau_m = 1$, the phase estimates are nearly 30 dB above the high SNR CRB (dashed lines).

To further investigate the problem of choosing the algorithm parameters τ_n and τ_m and the dependence of the selection rule on the dimensions of the observed field, we have repeated the foregoing Monte Carlo experiments for a much smaller observed field. In this set of experiments, only a 30×30 segment of the the field is observed. The Monte Carlo simulations were carried out for two different sets of τ_n and τ_m. In the first case, we chose $\tau_n = \tau_m = 1$ as in the first part of this example. In the second experiment, we chose $\tau_n = \tau_m = 4$. Note that high values of τ_n and τ_m cannot be used due to the small dimensions of the observed field. In these experiments, as well, the Monte Carlo simulations indicate that the estimator proposed in Section III-A yields unbiased

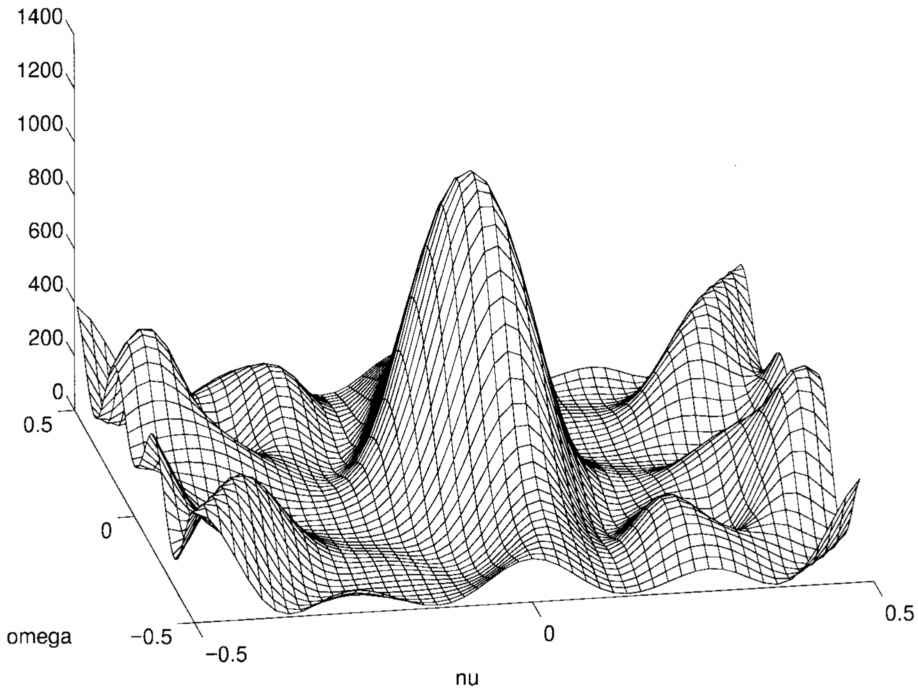


Fig. 4. Fourier transform of the MA field covariance function.

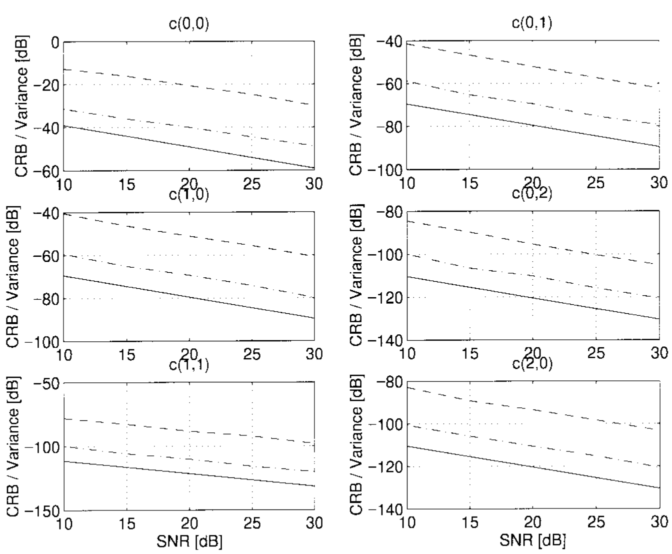


Fig. 5. Performance of the random amplitudepolynomial phase signal estimation algorithm for a 100×100 observed field. Solid lines denote the CRB assuming the SNR is high, dashed-dotted lines denote the experimental variance of the estimates for $\tau_n = \frac{N}{2}$ and $\tau_m = \frac{M}{2}$, while dashed lines denote the experimental variance of the estimates for $\tau_n = \tau_m = 1$.

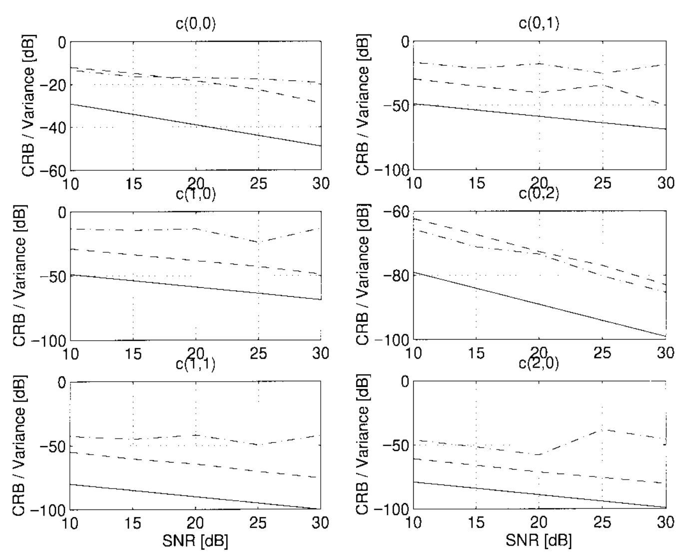


Fig. 6. Performance of the random amplitude polynomial phase signal estimation algorithm for a 30×30 observed field. Solid lines denote the CRB, assuming the SNR is high, dashed-dotted lines denote the experimental variance of the estimates for $\tau_n = \tau_m = 4$, while dashed lines denote the experimental variance of the estimates for $\tau_n = \tau_m = 1$.

estimates of the phase parameters, as the experimental bias is considerably smaller than the experimental standard deviation [except for the estimates of $c(0,0)$, which are slightly biased]. A comparison of the Monte Carlo results with the CRB that is computed assuming a high SNR is depicted in Fig. 6. The experimental results indicate that for a small observed field, better estimation results are obtained by choosing low values for τ_n and τ_m, contrary to the situation when the dimensions of the observed field are large.

Analysis of the performance of the proposed algorithm is beyond the scope of this paper. Such an analysis would provide an answer to the question of how to choose the algorithm parameters τ_n and τ_m as a function of the statistical properties of the amplitude field, the phase function, and the dimensions of the observed field. (Refer to [11] and [18] for detailed analyzes of this problem in the cases of 2-D and 1-D *constant* amplitude polynomial phase signals, respectively). Based on Theorem 1, and the derivation of the CRB, it is clear however, that while the high SNR CRB for the phase parameters is a function of SNR only, the performance of the proposed algorithm is a function of the power of the high-order moments of the field. Hence, given two amplitude fields of identical power, (and, hence, identical high SNR CRB in the Gaussian case), the performance of the algorithm when the amplitude is a zero mean field would be inferior to its performance when the amplitude field is positive. In the case of a Gaussian amplitude, the performance of the algorithm is strongly related to the rate of decay of the field autocorrelation function since fast decay of this function will enforce the choice of low values for τ_n and τ_m. Our experimental results indicate that such a choice will lead to less accurate estimates of the phase parameters.

VIII. Conclusion

In this paper, we presented a simple-to-implement and computationally efficient estimation algorithm for the parameters of 2-D signals with random amplitude and polynomial phase. The algorithm is based on the properties of the mean phase difference operator, which is introduced and analyzed. Assuming that the signal is observed in additive white Gaussian noise and that the amplitude field is Gaussian as well, we derived the Cramér–Rao lower bound on the error variance in jointly estimating the model parameters.

The performance of the algorithm in the presence of additive white Gaussian noise is illustrated by numerical examples and compared with the Cramér–Rao bound. In cases where the high-order moments of the amplitude field are not decaying too rapidly, the parameter estimates are shown to be unbiased, and the estimation error variance is shown to be close to the Cramér–Rao bound. From the examples shown, we conclude that the proposed phase estimation algorithm is quite robust in the presence of phase aliasing due to both low sampling rates and noise, as long as the true phase function is a continuous function of the coordinates. Since the phase model is inherently smooth, the proposed algorithm is not affected by the 2π ambiguities of the phase function.

References

[1] S. M. Song, S. Napel, N. J. Pelc, and G. H. Glover, "Phase unwrapping of MR phase images using Poisson equation," *IEEE Trans. Image Processing*, vol. 4, pp. 667–676, 1995.

[2] L. Axel and D. Morton, "Correction of phase unwrapping in magnetic resonance imaging," *Med. Phys.*, vol. 16, pp. 284–287, 1989.

[3] M. Hedley and D. Rosenfeld, "A new two-dimensional phase unwrapping algorithm for MRI images," *Magn. Reson. Med.*, vol. 24, pp. 177–181, 1992.

[4] T. R. Judge and P. J. Bryanston-Cross, "A review of phase unwrapping techniques in fringe analysis," *Opt. Lasers Eng.*, vol. 21, pp. 199–239, 1994.

[5] L. C. Graham, "Synthetic interferometer radar for topographic mapping," *Proc. IEEE*, vol. 62, pp. 763–768, 1974.

[6] Q. Lin, J. F. Vesecky, and H. A. Zebker, "New approaches in interferometric SAR data processing," *IEEE Trans. Geosci., Remote Sensing*, vol. 30, pp. 560–567, 1992.

[7] S. Peleg and B. Porat, "Estimation and classification of polynomial-phase signals," *IEEE Trans. Inform. Theory*, vol. 37, pp. 422–430, 1991.

[8] B. Friedlander and J. M. Francos, "An estimation algorithm for 2-D polynomial phase signals," *IEEE Trans. Image Processing*, vol. 5, pp. 1084–1087, June 1996.

[9] J. M. Francos and B. Friedlander, "Two-dimensional polynomial phase signals: Parameter estimation and bounds," *Multidimen. Syst. Signal Process.*, vol. 9, no. 2, pp. 173–205, 1998.

[10] S. Shamsundar, G. B. Giannakis, and B. Friedlander, "Estimating random amplitude polynomial phase signals: A cyclostationary approach," *IEEE Trans. Signal Processing*, vol. 43, pp. 492–505, 1995.

[11] J. M. Francos and B. Friedlander, "Optimal parameter selection in the phase differencing algorithm for 2-D phase estimation and unwrapping," *IEEE Trans. Signal Processing*, accepted for publication.

[12] U. Spagnolini, "2-D phase unwrapping and instantaneous frequency estimation," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 579–589, 1995.

[13] B. J. Super and A. C. Bovik, "Shape from texture using local spectral moments," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 333–343, 1995.

[14] T. L. Marzetta, "Two-dimensional linear prediction: Autocorrelation arrays, minimum-phase prediction error filters and reflection coefficient arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 725–733, 1980.

[15] B. Porat and B. Friedlander, "Computation of the exact information matrix of Gaussian time series with stationary random components," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 118–130, Jan. 1986.

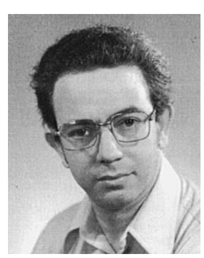
[16] J. M. Francos and B. Friedlander, "Parameter estimation of two-dimensional moving average random fields," *IEEE Trans. Signal Processing*, vol. 46, pp. 2157–2165, Aug. 1998.

[17] C. R. Rao, *Linear Statistical Inference and Its Applications*. New York: Wiley, 1965.

[18] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

Joseph M. Francos was born on November 6, 1959 in Tel-Aviv, Israel. He received the B.Sc. degree in computer engineering in 1982 and the D.Sc. degree in electrical engineering in 1990, both from the Technion—Israel Institute of Technology, Haifa.

From 1982 to 1987, he was with the Signal Corps Research Laboratories, Israeli Defense Forces. From 1991 to 1992, he was with the Department of Electrical Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, as a Visiting Assistant Professor. He then was with the Media Laboratory, Massachusetts Institute of Technology, Cambridge, during the summer of 1992. During 1993, he was with Signal Processing Technology, Palo Alto, CA. In 1993, he joined the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, Israel, where he is now a Lecturer. His current research interests are in parametric modeling and estimation of 2-D random fields, random fields theory, parametric modeling and estimation of nonstationary signals, image modeling, and texture analysis and synthesis.



Benjamin Friedlander (S'74–M'76–SM'82–F'87) received the B.Sc. and the M.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 1968 and 1972, respectively, and the Ph.D. degree in electrical engineering and the M.Sc. degree in statistics from Stanford University Stanford, CA, in 1976.

From 1976 to 1985, he was with Systems Control Technology, Inc., Palo Alto, CA. From November 1985 to July 1988, he was with Saxpy Computer Corporation, Sunnyvale, CA. Currently, he is with the University of California, Davis, and Signal Processing Technology, Ltd., Palo Alto.

Dr. Friedlander is the recipient of the 1983 ASSP Senior Award, the 1985 Award for the Best Paper of the Year from the European Association for Signal Processing (EURASIP), and the 1989 Technical Achievement Award of the Signal Processing Society.