©2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Analysis of Multicomponent Polynomial Phase Signals

Duc Son Pham, *Member, IEEE*, and Abdelhak M. Zoubir, *Senior Member, IEEE*

***Abstract*—While the theory of estimation of monocomponent polynomial phase signals is well established, the theoretical and methodical treatment of multicomponent polynomial phase signals (mc-PPSs) is limited. In this paper, we investigate several aspects of parameter estimation for mc-PPSs and derive the Cramér–Rao bound. We show the limits of existing techniques and then propose a nonlinear least squares (NLS) approach. We also motivate the use the Nelder–Mead simplex algorithm for minimizing the nonlinear cost function. The slight increase in computational complexity is a tradeoff for improved mean square error performance, which is evidenced by simulation results.**

***Index Terms*—High ambiguity function (HAF), Nelder–Mead algorithm, nonlinear least squares, nonstationary, polynomial phase signals.**

I. Introduction

In a large number of signal processing applications such as those in synthetic aperture radar (SAR) or in radio communications where the phase is continuously modulated, the signal of interest is *nonstationary*. The most common model used in parametric analysis of analytic and nonstationary signals is the polynomial phase signal (PPS) model, which is motivated by Weierstrass' theorem. This theorem implies that for a finite duration of observations, an arbitrary time-varying phase can be well approximated by a polynomial of sufficient order.

Results in the literature on estimation of PPS parameters are limited to the monocomponent case such as fast maximum likelihood (ML) [1], fast instantaneous frequency (IF) estimation [18], the dechirping algorithm [7], the nonlinear instantaneous least squares (NILS) [3], the well-known high ambiguity function (HAF) [19], [21]–[24], and its numerous extensions to time-varying and random amplitudes [8], [10], [11], [15], [16], [29], [31]. Other issues relating to monocomponent polynomial phase signal (mono-PPS) analysis include aliasing [2], nonuniform sampling schemes [13] and bootstrap model selection [32].

Manuscript received May 25, 2005; accepted January 17, 2006. This work was supported in part by the Australian Research Coucil under Grant DP0211849 while A. M. Zoubir was with the Curtin University of Technology. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Carlos H. Muravchik.

D. S. Pham was with the Communications and Signal Processing Group, Department of Electrical and Computer Engineering, and is now with the Institute for Multi-Sensor Processing and Content Analysis, Department of Computing, Curtin University of Technology, WA 6845 Perth, Australia (e-mail: dspham@ieee.org).

A. M. Zoubir is with the Signal Processing Group, Institute of Telecommunications, Darmstadt University of Technology, D-64283 Darmstadt, Germany (e-mail: zoubir@ieee.org).

Digital Object Identifier 10.1109/TSP.2006.882085

However, in a number of practical situations such as analysis of a nonstationary signal in the presence of another nonstationary jamming signal, the multicomponent model is more relevant.

While some theoretical aspects of mc-PPSs have been investigated (see [8] and [9]), practical algorithms for mc-PPSs seem to be absent from the literature. The main mc-PPS problems are twofold. First, the interactions between the components, which are often called *cross-terms,* give rise to undesired sinusoids in the high-order instantaneous moment (HIM), which strongly affects algorithms based on frequency estimation. Secondly, the principle of *demodulation* of monocomponent PPSs no longer works with mc-PPSs. In [20], an extension of the monocomponent case was presented for mc-PPSs. However, this does not solve the two main issues mentioned above. In [3], outlines of an extension of the NILS algorithm to the multicomponent case were given, but this method is computationally inefficient and the optimal choice of the parameters is unknown. Motivated by the Wigner–Hough transform, the integrated generalized ambiguity function was introduced in [5] to generalize the method in [4]. This method estimates two successive phase order coefficients at once, as opposed to conventional HAF-based methods. However, the algorithm is computationally demanding due to the need to form a two-dimensional grid search for every pair of phase coefficients. The most recent result for mc-PPSs is the product high-order ambiguity function (PHAF) introduced in [6] and [27]. This method exploits the property of the multilag HIM, that is, the autoterms are complex sinusoids with frequencies proportional to the product of lags. This property is not seen by any other cross-term. Hence, using multiple sets of lags and proper scaling, the autoterms can be enhanced which helps improve the identification of highest order polynomial phase coefficients. This method addressed the first issue in mc-PPS analysis, that is, cross-terms suppression. However, the estimation of the lower order phase coefficients and the amplitudes has not been successfully achieved. Besides, the analysis was given only for the monocomponent case and the technique in the original form can produce noticeable errors even in a noise-free condition.

In this paper, we quantitatively analyze the weakness of existing techniques when they are applied to mc-PPSs. Then we propose an NLS approach for estimating mc-PPSs which we believe is the most viable approach. When the noise is complex circular white Gaussian, the NLS estimates are also ML. Note that an ML method was also introduced in [9]. Our work differs from that in several aspects. First, we introduce an improved initialization technique which is more flexible and robust than the original HAF-based technique. Secondly, we propose

1053-587X/$20.00 © 2006 IEEE

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

to use Nelder–Mead's simplex algorithm to search for the solution of the highly nonlinear augmented NLS cost function. This algorithm is more robust than the quasi-Newton method used in [9]. Thirdly, we consider practical aspects such as implementation, numerical accuracy, and complexity of the algorithms. The numerical examples given here are only for second- and third-order mc-PPSs. However, extensions to higher orders are straightforward.

II. Signal Model and the Cramér Rao Bound

Consider a complex-valued K-component polynomial phase signal embedded in complex circular white Gaussian noise with variance σ^2

$$\begin{aligned} y(t) &= \sum_{k=1}^{K} s_k(t) + \nu(t) \\ &= \sum_{k=1}^{K} \alpha_k \exp\left\{ j \sum_{m=0}^{M} \omega_m^k t^m \right\} + \nu(t), \quad t \in \mathbb{R}. \end{aligned} \tag{1}$$

The problem is as follows: given N samples of the received signal $y(t)$ at time instances $t_0, \ldots, t_{N-1}$, find estimates of the amplitudes α_k and the polynomial phase coefficients ω_m^k for $k = 1, \ldots, K$, $m = 0, 1, \ldots, M$. We assume that all PPS components have the same order M and that both M and K are known due to the physical modelling of the problem, otherwise a model selection procedure similar to [32] can be developed. The results given subsequently can also be easily altered to cater for the more general case.

Denote $\boldsymbol{\omega}_k = [\omega_0^k, \ldots, \omega_M^k]^T$, $\mathbf{t} = [t_0, \ldots, t_{N-1}]^T$, $\mathbf{y} = [y(t_0), \ldots, y(t_{N-1})]^T$, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]^T$, $\boldsymbol{\nu} = [\nu(t_0), \ldots, \nu(t_{N-1})]^T$, $s_k(t, \boldsymbol{\omega}_k) = \exp\{j \sum_{m=0}^{M} \omega_m^k t^m\}$, $\mathbf{s}_k(\mathbf{t}, \boldsymbol{\omega}_k) = [s_k(t_0, \boldsymbol{\omega}_k), \ldots, s_k(t_{N-1}, \boldsymbol{\omega}_k)]^T = \mathbf{s}_k(\boldsymbol{\omega}_k)$, [1] $\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_K)$, and $\mathbf{S}(\boldsymbol{\omega}) = [\mathbf{s}_1(\boldsymbol{\omega}_1), \ldots, \mathbf{s}_K(\boldsymbol{\omega}_K)]$. We can rewrite (1) as

$$\begin{aligned} \mathbf{y} &= [\mathbf{s}_1(\mathbf{t}, \boldsymbol{\omega}_1), \ldots, \mathbf{s}_K(\mathbf{t}, \boldsymbol{\omega}_K)]\boldsymbol{\alpha} + \boldsymbol{\nu} \\ &= \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha} + \boldsymbol{\nu}. \end{aligned} \tag{2}$$

The objective is to estimate $\boldsymbol{\omega}$, $\boldsymbol{\alpha}$, and σ^2. Denote the unknown parameter set $\boldsymbol{\vartheta} = (\boldsymbol{\alpha}, \boldsymbol{\omega}, \sigma^2)$. The pdf of the noise is given by

$$p(\mathbf{y}; \boldsymbol{\vartheta}) = \frac{1}{(\pi\sigma^2)^N} \exp\left\{ -\frac{(\mathbf{y} - \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha})^H (\mathbf{y} - \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha})}{\sigma^2} \right\}.$$

The log-likelihood function is

$$\mathcal{L}(\boldsymbol{\vartheta}) = -N \ln \pi - N \ln \sigma^2 - \frac{1}{\sigma^2} (\mathbf{y} - \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha})^H (\mathbf{y} - \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha}). \tag{3}$$

The (i, j)th element of Fisher's information matrix is given by

$$J_{ij} = -\mathsf{E}\left\{ \frac{\partial^2 \mathcal{L}(\boldsymbol{\vartheta})}{\partial \vartheta_i \partial \vartheta_j} \right\}. \tag{4}$$

[1]The parameter $\mathbf{t}$ is omitted for notational simplification.

It can be shown that the structure of Fisher's information matrix is

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\omega}} & \mathbf{0} \\ \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\alpha}} & \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_{\sigma^2\sigma^2} \end{bmatrix}. \tag{5}$$

Recall that

$$J_{ij} = \mathrm{tr}\left\{ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \vartheta_i} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \vartheta_j} \right\} + 2\Re\left\{ \frac{\partial \boldsymbol{\mu}^H}{\partial \vartheta_i} \mathbf{R}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \vartheta_j} \right\} \tag{6}$$

where $\mathbf{R} = \sigma^2 \mathbf{I}$ is the covariance matrix of the noise and

$$\begin{aligned} \boldsymbol{\mu} &= \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha} \\ &= \sum_{k=1}^{K} \mathbf{s}_k(\boldsymbol{\omega}_k)\alpha_k. \end{aligned} \tag{7}$$

It can be easily shown that

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}}{\partial \alpha_k} &= \frac{\partial}{\partial \alpha_k} \sum_{k=1}^{K} \mathbf{s}_k(\boldsymbol{\omega}_k)\alpha_k \\ &= \mathbf{s}_k(\boldsymbol{\omega}_k). \end{aligned} \tag{8}$$

Therefore

$$\mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \frac{2}{\sigma^2} \Re\left\{ \mathbf{S}(\boldsymbol{\omega})^H \mathbf{S}(\boldsymbol{\omega}) \right\}. \tag{9}$$

Now, for the block $\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}$, we note that

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}}{\partial \omega_m^k} &= \frac{\partial}{\partial \omega_m^k} \mathbf{s}_k(\mathbf{t}, \boldsymbol{\omega}_k)\alpha_k \\ &= \frac{\partial}{\partial \omega_m^k} \begin{bmatrix} \alpha_k \exp\left\{ j \sum_{m'=0}^{M} \omega_{m'}^k t_0^{m'} \right\} \\ \cdots\cdots \\ \alpha_k \exp\left\{ j \sum_{m'=0}^{M} \omega_{m'}^k t_{N-1}^{m'} \right\} \end{bmatrix} \\ &= j\alpha_k \mathbf{T}^m \mathbf{s}_k(\boldsymbol{\omega}_k) \end{aligned} \tag{10}$$

where $\mathbf{T} = \mathrm{diag}\{t_0, \ldots, t_{N-1}\}$. Hence, the element of Fisher's information matrix that is due to the interaction between ω_m^k and $\omega_{m'}^{k'}$ is

$$J_{\boldsymbol{\omega}\boldsymbol{\omega}}\left(\omega_m^k, \omega_{m'}^{k'}\right) = \frac{2}{\sigma^2} \Re\left\{ \alpha_k \alpha_{k'} \mathbf{s}_k(\boldsymbol{\omega}_k)^H \mathbf{T}^{m+m'} \mathbf{s}_{k'}(\boldsymbol{\omega}_{k'}) \right\}. \tag{11}$$

Using the same notation, it follows from (8) and (11) that

$$J_{\boldsymbol{\omega}\boldsymbol{\alpha}}\left(\omega_m^k, \alpha_{k'}\right) = \frac{2}{\sigma^2} \Im\left\{ \alpha_k \mathbf{s}_k(\boldsymbol{\omega}_k) \mathbf{T}^m \mathbf{s}_{k'}(\boldsymbol{\omega}_{k'}) \right\}. \tag{12}$$

From the likelihood function (3), it can be shown that

$$\mathbf{J}_{\sigma^2\sigma^2} = N/\sigma^4. \tag{13}$$

The Cramér–Rao bound on the covariance of the estimate is found by taking the inverse of Fisher's information matrix $\mathbf{C} = \mathbf{J}^{-1}$.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

There are some observations on the Cramér–Rao bound for mc-PPSs.

- Compared with the results for monocomponent PPSs [22], [23], the Cramér–Rao bound for mc-PPSs also depends on the parameters themselves. This suggests that, depending on the values of the parameters, the bound of each parameter may change significantly, such as when there are considerable crossings between components.
- Since the Cramér–Rao bound depends on the parameters, it is likely that the SNR thresholds of nonlinear techniques for mc-PPSs will depend on each specific case, which is different from many known results for the monocomponent case.
- The Cramér–Rao bound is fully specified and can be numerically evaluated.

The two component case is of theoretical and practical interest and can be found, for example, in the removal of a nonstationary jamming signal mentioned above. For this case, some special properties of the Cramér–Rao bound are observed in the following results.

Corollary 1: Consider the mc-PPS model (1). When $K = 2$:

- the Cramér-Rao bounds on the estimates of the amplitudes α_1 and α_2, i.e., the diagonal elements of $\mathbf{C}$ that correspond to α_1 and α_2 are the same

$$\mathrm{CRB}\{\alpha_1\} = \mathrm{CRB}\{\alpha_2\} \tag{14}$$

  regardless of the values of all other parameters.
- furthermore, if the amplitudes of the two components are the same, i.e., $\alpha_1 = \alpha_2 = \alpha$, the Cramér–Rao bounds of the polynomial phase parameters of the two components are the same, i.e.,

$$\mathrm{CRB}\left\{\omega_m^1\right\} = \mathrm{CRB}\left\{\omega_m^2\right\}, \quad m = 0, 1 \ldots, M. \tag{15}$$

The proof of these results is given in Appendix A. The results suggest that the asymptotic accuracy of the amplitude estimate of a weaker component is the same as the other regardless of their powers and vice versa. If they have the same power, estimating one component is of the same difficulty as estimating the other.

III. Estimation Techniques for mc-PPSs

A. Product High-Order Ambiguity Function

The PHAF method [6] is a generalized version of the HAF method originally introduced by Peleg and Porat [21]. We define the multilag high-order instantaneous moment (ml-HIM) of $y(t)$ as follows:

$$y_1(t) = y(t) \tag{16}$$

$$y_2(t; \boldsymbol{\tau}_1) = y_1(t + \tau_1) y_1(t - \tau_1)^* \tag{17}$$

$$\ldots \tag{18}$$

$$y_M(t; \boldsymbol{\tau}_{M-1}) = y_{M-1}(t + \tau_{M-1}; \boldsymbol{\tau}_{M-1}) \times y_{M-1}(t - \tau_{M-1}; \boldsymbol{\tau}_{M-1})^* \tag{19}$$

where $\boldsymbol{\tau}_i = [\tau_1, \tau_2, \ldots, \tau_{i-1}]$. One can see that the original HIM introduced by Peleg and Porat is a special case of this definition where the lags are the same, i.e., $\tau_1 = \ldots = \tau_{M-1} = \tau$. The multilag HAF is defined as the finite Fourier transform of the ml-HIM

$$Y_m(f; \boldsymbol{\tau}_{M-1}) = \sum_{t=t_0}^{t=t_{N-1}} y_M(t; \boldsymbol{\tau}_{M-1}) e^{-j2\pi f t}. \tag{20}$$

In the case of a mono-PPS $y(t) = \alpha \exp\{j \sum_{m=0}^{M} \omega_m t^m\}$, it is true that the Mth-order HIM contains only a single sinusoid whose frequency is linearly proportional to the highest phase coefficient ω_M. Hence, one can perform peak picking over the discrete Fourier transform (DFT) of the HIM and estimate ω_M. The lower order phase coefficients can be estimated in a similar fashion based on the demodulated signal. For details, see, for example, [19].

In the case of a K component Mth-order PPS, one can show by direct evaluation that the ml-HIM contains K sinusoids that correspond to K autoterms, each having a frequency proportional to the highest order phase coefficient ω_M^k. However, the ml-HIM also contains a large number of cross-terms which are also Mth-order PPSs and spurious sinusoids. With the ml-HAF, the spectrum of the cross-terms spreads over all frequencies and can affect the peak picking procedure that we use with mono-PPSs. Except for second-order mc-PPSs in which the spectrum of the cross-terms spreads *evenly* over frequencies, it is known [26] that the spectrum of one particular cross-term of order M is the convolution of M spectra of orders from $m = 1 \rightarrow M$, and each has a nonlinear power behavior $f^{-((m-2)/(m-1))}$, $m \geq 2$. The cross-terms, therefore, can introduce location shift of the peaks or incorrect peak picking, leading to significant errors of the estimates.

To reduce the effects of cross-terms, the PHAF was proposed in [6] after observing that only the autoterms have the peculiar property that they are single sinusoids and their frequencies are proportional to ω_M^k and the product of lags. Hence, by using different sets of lags and after proper scaling and multiplications of scaled ml-HAFs, the autoterms are always enhanced more significantly than the cross-terms, which strengthen the discriminant ability. Mathematically, suppose that we have L sets of lags

$$\mathbf{T}_{M-1}^L = \left[\boldsymbol{\tau}_{M-1}^{(1)}, \boldsymbol{\tau}_{M-1}^{(2)}, \ldots, \boldsymbol{\tau}_{M-1}^{(L)}\right]$$
$$\boldsymbol{\tau}_{M-1}^{(l)} = \left[\tau_1^{(l)}, \tau_2^{(l)}, \ldots, \tau_{M-1}^{(l)}\right]^T. \tag{21}$$

The PHAF is defined as follows:

$$Y_M^L(f; \mathbf{T}_{M-1}^L) = \prod_{l=1}^{L} Y_M\left(\beta^{(l)} f; \boldsymbol{\tau}_{M-1}^{(l)}\right) \tag{22}$$

where the scale factor is

$$\beta^{(l)} = P_M\left(\boldsymbol{\tau}_{M-1}^{(l)}, \boldsymbol{\tau}_{M-1}^{(1)}\right) = \frac{\left(\prod_{k=1}^{M-1} \tau_k^{(l)}\right)}{\left(\prod_{k=1}^{M-1} \tau_k^{(1)}\right)}. \tag{23}$$

The introduction of the scale factor is to align the autoterms from different sets of lags. It is noted that the mechanism of the PHAF is analogous to integration along the curve of $Y_M^L(f; \mathbf{T}_{M-1}^L)$ in the ambiguity domain whose two axes are

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

$\mathbf{T}_{M-1}^{L}$ and f. Like many nonlinear techniques, the PHAF method is strongly affected by finite sample effects in which the noise-free error is caused by interference from the other components. In what follows, we give a quantitative error analysis of this method. For simplicity, consider a two-component second-order discrete-time PPS in the absence of noise with uniform sampling such that $t_n = n$

$$y(n) = \alpha_1 \exp\left\{j\left(\omega_0^1 + \omega_1^1 n + \omega_2^1 n^2\right)\right\} + \alpha_2 \exp\left\{j\left(\omega_0^2 + \omega_1^2 n + \omega_2^2 n^2\right)\right\}. \quad (24)$$

Suppose that the first component is of interest. We shall describe a perturbation analysis on the estimate of ω_2^1 using the PHAF method. Note that the PHAF is essentially similar to the HAF method. Lag diversity is to suppress spurious peaks. Here we perform perturbation analysis about the true parameter. Hence, it is reasonable to consider only one lag, which reduces to the HAF method indeed. The discrete version of the high-order instantaneous moment with a lag k is

$$y_2(n) = y(n+k)y(n-k)^*. \quad (25)$$

The HAF is

$$Y_2(\omega) = \sum_{n=0}^{N-1} y_2(n)\exp\{-j\omega n\}. \quad (26)$$

Let $\omega_0 = 4k\omega_2^1$. When there is only one component, i.e., $\alpha_2 = 0$, it is true that the peak of $|Y_2(\omega)|$ occurs at ω_0. However, in the presence of the second component, it can be shown (see Appendix B) that the displacement of the peak is approximately given by (34), where all the sums are taken from $n = 0$ to $N-1$, $p(n) = \Re\{y_2(n)\exp\{-j\omega_0\}\}$, and $q(n) = \Im\{y_2(n)\exp\{-j\omega_0\}\}$. Note that this error has nothing to do with the accuracy of the computer; it is the nature of the PHAF algorithm.

By using similar arguments, it can be shown that for higher orders and when there are more components, the error does exist. The error on PHAF estimates of the highest order coefficients will propagate to other lower order coefficients with standard techniques. However, this error propagation effect can be suppressed if one further performs a fine search. That motivates the nonlinear least squares (NLS) approach to be described next.

B. Nonlinear Least Squares

To overcome the fundamental problem with PHAF, we propose an NLS approach to estimating the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$ in (2)

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\omega}}) = \arg\min_{\boldsymbol{\alpha},\boldsymbol{\omega}} \left(\mathbf{y} - \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha}\right)^H \left(\mathbf{y} - \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha}\right). \quad (27)$$

To reduce the number of unknowns in (27), we further split the polynomial phase parameters into $\boldsymbol{\omega}_k = (\omega_0^k, \boldsymbol{\theta}_M)$, $k = 1, \ldots, K$, and introduce $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K)$. Essentially, $\boldsymbol{\theta}_k$ consists of M polynomial phase parameters with orders ranging from one to M. Introduce $\boldsymbol{\eta} = [\omega_0^1, \omega_0^2, \ldots, \omega_0^K]^T$ and decompose

$$\boldsymbol{S}(\boldsymbol{\omega})\boldsymbol{\alpha} = \mathbf{H}(\boldsymbol{\theta})\mathbf{c}(\boldsymbol{\eta}, \boldsymbol{\alpha}) \quad (28)$$

where

$$\mathbf{c}(\boldsymbol{\eta}, \boldsymbol{\alpha}) = \left[\alpha_1 \exp\left\{j\omega_0^1\right\}, \ldots, \alpha_K \exp\left\{j\omega_0^K\right\}\right]^T \in \mathbb{C}^K \quad (29)$$

$$\mathbf{H}(\boldsymbol{\theta}) = [\mathbf{h}_1(\boldsymbol{\theta}_1), \ldots, \mathbf{h}_K(\boldsymbol{\theta}_K)] \in \mathbb{C}^{N\times K} \quad (30)$$

and $\mathbf{h}_k(\boldsymbol{\theta}_k) \in \mathbb{C}^N$ is a column vector whose nth element is $\exp\{j\sum_{m=1}^{M}\omega_m^k t_{n-1}^m\}$. Note that, when the noise is complex circular white Gaussian and if the estimate of $\boldsymbol{\theta}$, which we denote by $\hat{\boldsymbol{\theta}}$, is available, then $\boldsymbol{\eta}$ and $\boldsymbol{\alpha}$ are estimated from

$$\hat{\mathbf{c}} = \mathbf{H}(\hat{\boldsymbol{\theta}})^{\dagger}\mathbf{y} \quad (31)$$

$$\hat{\boldsymbol{\alpha}} = |\hat{\mathbf{c}}| \quad (32)$$

$$\hat{\boldsymbol{\eta}} = \angle\{\hat{\mathbf{c}}\} \quad (33)$$

where $\mathbf{H}(\hat{\boldsymbol{\theta}})^{\dagger} = (\mathbf{H}(\hat{\boldsymbol{\theta}})^H\mathbf{H}(\hat{\boldsymbol{\theta}}))^{-1}\mathbf{H}(\hat{\boldsymbol{\theta}})^H$ denotes the pseudoinverse of $\mathbf{H}(\hat{\boldsymbol{\theta}})$. It then follows that the nonlinear optimization problem (27) can be solved more conveniently by considering the following equivalent optimization problem on the augmented NLS cost function:

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= \arg\min_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) \\ &= \arg\min_{\boldsymbol{\theta}} \left\{-\mathbf{y}^H\mathbf{P}_{\mathbf{H}(\boldsymbol{\theta})}\mathbf{y}\right\}\end{aligned} \quad (35)$$

where $\mathbf{P}_{\mathbf{H}(\boldsymbol{\theta})} = \mathbf{H}(\boldsymbol{\theta})(\mathbf{H}(\boldsymbol{\theta})^H\mathbf{H}(\boldsymbol{\theta}))^{-1}\mathbf{H}(\boldsymbol{\theta})^H$ denotes the projection matrix onto the subspace of $\mathbf{H}(\boldsymbol{\theta})$. Asymptotic theory of NLS estimation can be found in, for example, [30]. Note than when the noise is complex circular white Gaussian, the NLS estimates are ML; hence desirable properties such as unbiasedness and asymptotic efficiency are attained. In [3], the so-called NILS was proposed. The intuitive idea gives very good results but suffers the heuristic choice of parameters for it to perform well. Also, the examples shown in [3] are limited to two parameters of a mono-PPS. For mc-PPSs, such conclusions cannot be made directly. Besides, statistical properties of the NILS are unknown and an additional computational cost is introduced.

As with any nonlinear optimization problem, the major issues of (35) are initialization and a robust numerical algorithm for finding its solution. In what follows, we discuss initialization and a numerical approach for solving the nonlinear optimization (35).

$$\delta\omega \approx \frac{\left(\sum np(n)\right)\left(\sum q(n)\right) - \left(\sum nq(n)\right)\left(\sum p(n)\right)}{\left(\sum nq(n)\right)^2 + \left(\sum np(n)\right)^2 - \left(\sum n^2p(n)\right)\left(\sum p(n)\right)\left(\sum n^2q(n)\right)\left(\sum q(n)\right)} \quad (34)$$

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

TABLE I
EXAMPLE OF A SECOND ORDER mc-PPS

	True values		PHAF estimates	
	Component 1	Component 2	Component 1	Component 2
α	1	1	0.99821699938630	0.99821699938624
ω_0	0	0	-0.07853902562474	0.07853944287279
ω_1	2.19911485751286	1.25663706143592	2.20308214445873	1.25266976100833
ω_2	0.00490873852123	-0.00490873852123	0.00487309553696	-0.00487309553697

1) Initialization: The PHAF method can be used to initialize the estimates with some modifications. After estimating the highest order phase coefficients ω_M^k, $k = 1, \ldots, K$, we adopt the demodulation strategy to obtain the lower order phase coefficients for each component. Note that due to the augmented NLS cost function (35), we do not need to estimate the amplitudes and initial phases. It can be seen that, due to the presence of the other components, the demodulated sequence for component k at order m also contains $K-1$ components of order M. It can be inferred that except for the case $M = 2$, where the contribution from the other components spreads approximately evenly over the spectrum, the estimates of lower order phase coefficients based on Fourier analysis of the HIM of the demodulated sequence are subject to incorrect peak picking.

2) Numerical Algorithm: There are numerical algorithms available to solve the unconstrained optimization problem (35) (see, for example, [14] and [25]). In [9], the quasi-Newton method was used and the derivatives of the cost function were analytically derived. However, since the problem is highly nonlinear, it has been shown in practice that this approach is not robust and there is a great chance that the estimates get trapped in an undesirable local minimum. To illustrate this, consider the following example: a two-component second-order PPS with the same amplitudes in a noise-free environment. The parameters and their initial PHAF estimates are given in Table I. Fig. 1 plots the value of the cost function at each iteration for both algorithms. Because there is no noise, one would expect the value of the cost function to be zero at the true values. It can be seen from Fig. 1 that the quasi-Newton algorithm seems to be trapped in a local minimum. On the other hand, the Nelder–Mead algorithm attains better computer accuracy.

In what follows, we propose to use the robust Nelder–Mead simplex (NM) algorithm [12], [17] to find the NLS solution. The augmented NLS cost function $f(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \mathbb{R}^n$ is to be minimized with a starting point $\boldsymbol{\theta}_0$. The kth iteration of the NM algorithm for solving this problem is summarized in Table II, where the parameters are reflection $\rho = 1$, expansion $\chi = 2$, contraction $\gamma = 1/2$, and shrinkage $\sigma = 1/2$ [12]. One can see that if the initial simplex encloses the convergence region, then it is likely that the algorithm will reach the global solution. For strictly convex problems, the quasi-Newton method may be faster. However, for general problems, the NM algorithm is more robust and suitable, especially when these are highly nonlinear. The simplex algorithm is more suitable for a small number of parameters rather than a large one.

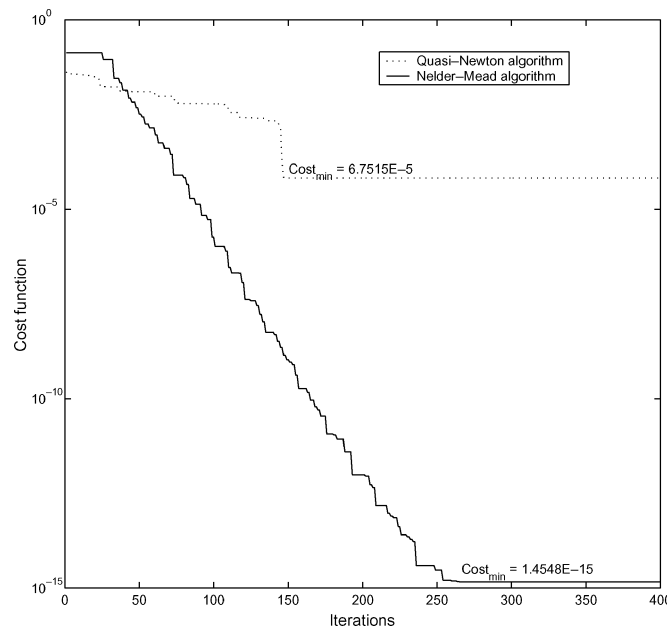

Fig. 1. Comparison between the quasi-Newton and Nelder–Mead algorithms for a second-order mc-PPS.

IV. SIMULATION RESULTS

In the first example, we consider a two-component second-order PPS embedded in complex circular white Gaussian noise. The parameters of the signal are $\alpha_1 = \alpha_2 = 1$, $(\omega_0^1, \omega_1^1, \omega_2^1) = (0, 20\pi, -0.22\pi/N)$, $(\omega_0^2, \omega_1^2, \omega_2^2) = (0, 0.80\pi, -0.31\pi/N)$, where N is the number of observations and the sampling interval is $\Delta = 1$. Suppose component 1 is of interest. It is noted that for these settings, the instantaneous frequencies of the two components cross each other (see Fig. 2). For the PHAF, the set of lags are (16,14,13,11), in which the first lag is optimized in the HAF sense. Note that for a second order mc-PPS, we need only one lag for each set. First, we fix the number of observations $N = 64$ and vary the SNR from -4 to 6 dB and measure the MSEs of the parameters of component 1 obtained from each method. The Cramér–Rao bound is numerically obtained and included in the plots. The MSE performance for ω_2^1 is depicted in Fig. 2, where the following can be seen.

- The PHAF method exhibits a strong finite sample effect, i.e., there is an SNR point (2 dB) where the MSE cannot be decreased further. This agrees with the analysis given earlier.
- The proposed NLS approach achieves a smaller MSE than the PHAF method. It has an SNR threshold about 3 dB.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

TABLE II
SUMMARY OF THE NELDER–MEAD SIMPLEX ALGORITHM

- **Step 1 (Order).** Sort and label the vertices $\boldsymbol{\theta}_1^{(k)}, \ldots, \boldsymbol{\theta}_{n+1}^{(k)}$ that constitute a simplex such that
$$f(\boldsymbol{\theta}_1^{(k)}) \leq f(\boldsymbol{\theta}_2^{(k)}) \ldots \leq f(\boldsymbol{\theta}_{n+1}^{(k)}).$$
For the first iteration, the other n vertices can be chosen by perturbing $\boldsymbol{\theta}_0$ in each dimension: $\boldsymbol{\theta}_0 + \epsilon \mathbf{e}_i, i = 1, \ldots, n$, where $\epsilon \ll 1$.
- **Step 2 (Reflect).** Compute the reflection point $\boldsymbol{\theta}_r^{(k)} = (1+\rho)\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\theta}_n^{(k)} - \rho \boldsymbol{\theta}_{n+1}^{(k)}$. If $f(\boldsymbol{\theta}_1^{(k)}) \leq f(\boldsymbol{\theta}_r^{(k)}) < f(\boldsymbol{\theta}_n^{(k)})$, the reflection point $\boldsymbol{\theta}_r^{(k)}$ is accepted and proceed to the next iteration. Otherwise, if $f(\boldsymbol{\theta}_1^{(k)}) > f(\boldsymbol{\theta}_r^{(k)})$ then go to Step 3. If $f(\boldsymbol{\theta}_n^{(k)}) \leq f(\boldsymbol{\theta}_r^{(k)})$ go to Step 4.
- **Step 3 (Expand).** Compute the expansion point $\boldsymbol{\theta}_e^{(k)} = (1+\rho\chi)\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\theta}_n^{(k)} - \rho\chi \boldsymbol{\theta}_{n+1}^{(k)}$. If $f(\boldsymbol{\theta}_e^{(k)}) < f(\boldsymbol{\theta}_r^{(k)})$ then accept $\boldsymbol{\theta}_e^{(k)}$, otherwise accept $\boldsymbol{\theta}_r$ and go to the next iteration.
- **Step 4 (Contract).** If $f(\boldsymbol{\theta}_n^{(k)}) \leq f(\boldsymbol{\theta}_r^{(k)}) < f(\boldsymbol{\theta}_{n+1}^{(k)})$ then perform an *outside contraction*. If $f(\boldsymbol{\theta}_r^{(k)}) > f(\boldsymbol{\theta}_{n+1}^{(k)})$ then perform an *inside contraction*.
  - *Outside*. The *outside contraction* point is computed as follow. $\boldsymbol{\theta}_c^{(k)} = (1+\rho\gamma)\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\theta}_n^{(k)} - \rho\gamma \boldsymbol{\theta}_{n+1}^{(k)}$. If $f(\boldsymbol{\theta}_c^{(k)}) \leq f(\boldsymbol{\theta}_r^{(k)})$ then accept $\boldsymbol{\theta}_c$, otherwise go to Step 5.
  - *Inside*. The *inside contraction* point is computed as follows: $\boldsymbol{\theta}_{cc}^{(k)} = (1-\gamma)\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\theta}_n^{(k)} + \gamma \boldsymbol{\theta}_{n+1}^{(k)}$. If $f(\boldsymbol{\theta}_c^{(k)}) \leq f(\boldsymbol{\theta}_{n=1}^{(k)})$ then accept $\boldsymbol{\theta}_{cc}$, otherwise go to Step 5.
- **Step 5 (Shrink).** Compute another extra n vertices $\mathbf{v}_i^{(k)} = \boldsymbol{\theta}_1^{(k)} + \sigma(\boldsymbol{\theta}_i^{(k)} - \boldsymbol{\theta}_1^{(k)}),\ i = 2, \ldots, n+1$. These $2n+1$ vertices will be sorted and selected in the next iteration.

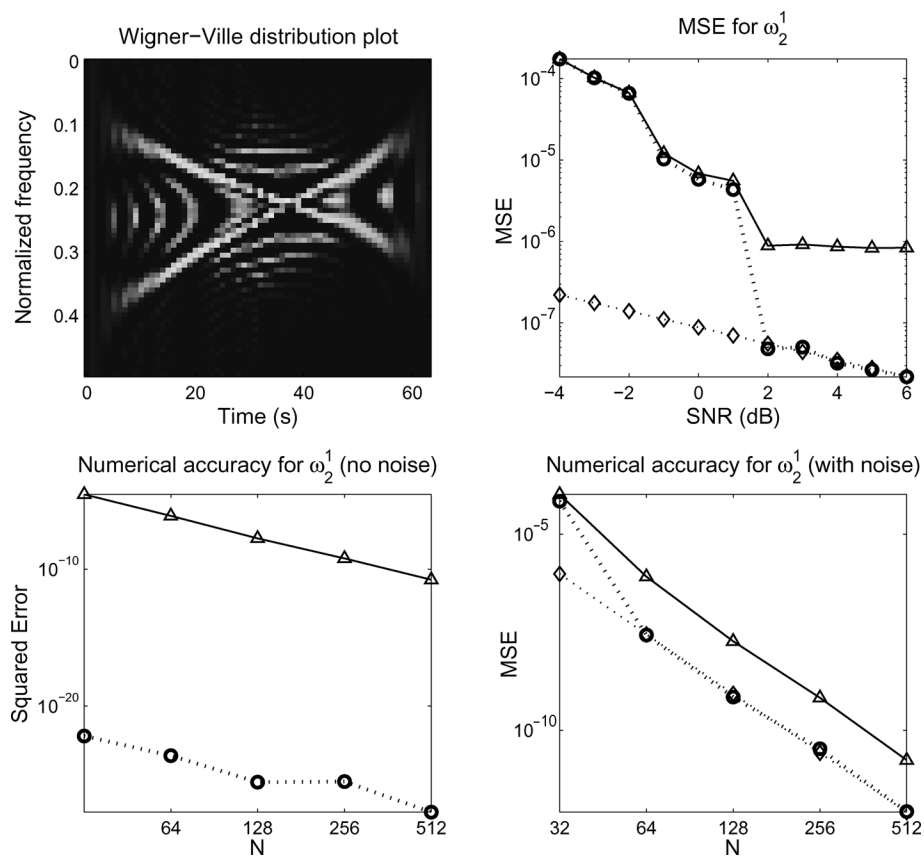

Fig. 2. Example of a second-order mc-PPS. Top left: the Wigner-Ville distribution. Top right: MSE performance for ω_2^1. Bottom left: numerical accuracy for ω_2^1 with respect to N in the absence of noise. Bottom right: numerical accuracy for ω_2^1 with respect to N with noise at SNR $= 5$ dB. Legend: PHAF $(-\triangle-)$, NLS $(\cdots \bigcirc \cdots)$, Cramér–Rao bound $(\cdots \diamond \cdots)$.

- In the simulation settings, we selected maximum computer accuracy (i.e., the tolerance is 2.204×10^{-16}). It is found that, on average, the convergence rate of the NLS is approximately 30%. The number of function evaluations is about 140 for this accuracy.

With the same settings, we also investigated the MSE performance of the estimators when N varies significantly. In the first scenario, we consider the mc-PPS without noise. The purpose of this study is to analyze the variation of finite sample errors as N becomes large. The results are given in Fig. 2 (bottom left). As can be seen, while the NLS maintains the maximum computer accuracy, the errors of the PHAF are relatively considerable when N varies from 32 to 512. In the second scenario, we maintain an SNR $= 5$ dB and investigate the MSE performance

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

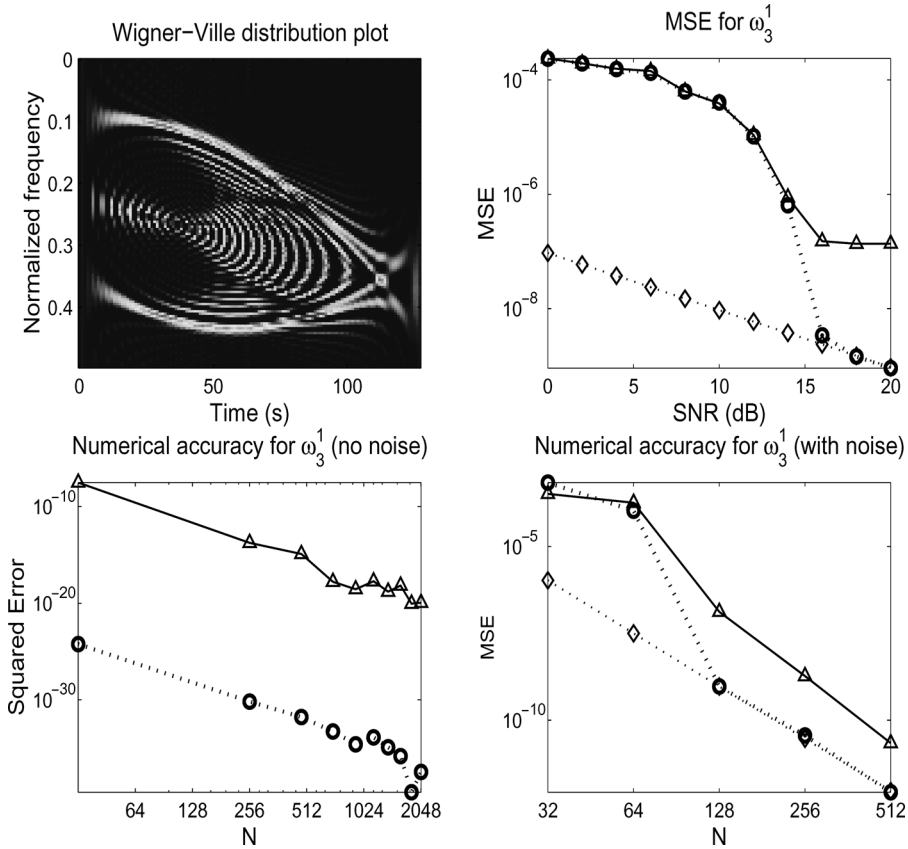

Fig. 3. Example of a third-order mc-PPS. Top left: the Wigner-Ville distribution. Top right: MSE performance for ω_3^1. Bottom left: numerical accuracy for ω_3^1 with respect to N in the absence of noise. Bottom right: numerical accuracy for ω_3^1 with respect to N with noise at SNR $= 20$ dB. Legend: PHAF $(-\triangle-)$, NLS $(\cdots\bigcirc\cdots)$, Cramér–Rao bound $(\cdots\diamond\cdots)$.

when N varies from 32 to 512. The results are depicted in Fig. 2 (bottom right). The MSEs mostly decrease with N. The sample threshold for the NLS method is less than 100, while the PHAF maintains a considerable loss from the Cramér–Rao bound over the simulated range of N. This can be most relevantly explained by the finite sample effect due to interference from the other components.

Next, we consider a two-component third-order mc-PPS embedded in complex circular white Gaussian noise. The parameters of the signal are $\alpha_1 = \alpha_2 = 1$, $(\omega_0^1,\ \omega_1^1,\ \omega_2^1,\ \omega_3^1) = (0,\ 20\pi,\ -0.10\pi/N,\ 0.30\pi/N^2)$, $(\omega_0^2,\ \omega_1^2,\ \omega_2^2,\ \omega_3^2) = (0, 0.70\pi,\ 0.40\pi/N,\ -0.30\pi/N^2)$, $N = 128$. For these settings, the instantaneous frequencies of the two components cross each other (see Fig. 3). The MSE performance is shown in Fig. 3. From this plot, the following observations are noted.

- When the SNR is small, NLS does not improve over the PHAF (except for the estimate of ω_0^1). The SNR threshold for these particular problem settings is 16 dB, which is more significant than the example for a second-order mc-PPS. This is best explained by the nonlinearities introduced by higher order mc-PPS to the optimization problem.
- After a 16 dB point where NLS attains the Cramér–Rao bound, the PHAF exhibits finite sample error, which was also observed in the previous example.
- In this simulation, we set the same tolerance as in the previous example. With NLS, it is found that, on average, the convergence rate of NLS is approximately 20%, which is less than the one in the other example. The average number of function evaluations is about 200 for this accuracy.

Again, we also investigate the MSE performance when N varies from 32 to 512. The results are plotted in Fig. 3 in the absence of noise. As can be seen, both the PHAF and the NLS improve the square error performance with large values of N. However, the gap between the PHAF and the NLS remains nearly the same. Intensive simulations beyond this range also confirm this. Results when noise with an SNR $= 20$ dB are indicated in Fig. 3. It is shown that NLS significantly improves over the PHAF method at large values of N. This implies that NLS only performs well when both the SNR and N are larger than their respective thresholds.

Based on the observations for second- and third-order mc-PPSs, we can anticipate that for fourth and higher orders, the SNR and N thresholds will be much higher. Since the optimization problem is extremely nonlinear, it is likely that convergence and numerical accuracy can only be satisfactory at high SNR and large N.

V. Conclusion

We have presented an analytical study of mc-PPSs embedded in complex circular white Gaussian noise. The Cramér–Rao bound was derived and found to be dependent on the parameters of the underlying mc-PPS. We have shown that the PHAFmethod suffers from a considerable finite sample error

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

due to its nonlinearity when applied to multicomponent signals, and this has been convincingly demonstrated via simulation results. We have also proposed an NLS approach to estimate the parameters of mc-PPSs, which overcomes the limitations of existing techniques, with some increase in computational complexity. The Nelder–Mead algorithm is suggested to be used with the NLS approach so as to achieve better performance, especially at a high SNR and a large sample size. The performance of the proposed NLS approach has been intensively investigated for several cases and the numerical properties have been identified.

APPENDIX A
PROOF OF COROLLARY 1

Proof: First, we prove the first part of the corollary. For simplicity, the following assumptions and shorthand notations are used.

- Assume $(2/\sigma^2) = 1$ so that the common factor of Fisher's information matrix can be ignored.
- Use $\mathbf{s}_1$ for $\mathbf{s}_1(\boldsymbol{\omega}_1)$ and $\mathbf{s}_2$ for $\mathbf{s}_2(\boldsymbol{\omega}_2)$.
- Ignore $\mathbf{J}_{\sigma^2\sigma^2}$, as it does not affect subsequent results. Hence, Fisher's information matrix now becomes

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\omega}} \\ \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\alpha}} & \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} \end{bmatrix}. \tag{36}$$

The Cramér–Rao bound is therefore

$$\mathbf{C} = \mathbf{J}^{-1} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\alpha}\boldsymbol{\omega}} \\ \mathbf{C}_{\boldsymbol{\omega}\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\omega}\boldsymbol{\omega}} \end{bmatrix} \tag{37}$$

where, according to the partitioned matrix inverse theorem [28, p. 53]

$$\mathbf{C}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \left(\mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} - \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\omega}}\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}^{-1}\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\alpha}}\right)^{-1}. \tag{38}$$

It follows from (9) that

$$\mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \begin{bmatrix} N & \Re\left\{\mathbf{s}_1^H\mathbf{s}_2\right\} \\ \Re\left\{\mathbf{s}_1^H\mathbf{s}_2\right\} & N \end{bmatrix}. \tag{39}$$

Denote

$$\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\alpha}} = [\mathbf{m}_1\ \mathbf{m}_2]. \tag{40}$$

Using the fact that $\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}$ is symmetric positive definite and from (38) and (39), it follows from the partitioned matrix inverse theorem that a necessary and sufficient condition for the diagonal elements of $\mathbf{C}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}$ to be the same is

$$\mathbf{m}_1^T\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}^{-1}\mathbf{m}_1 = \mathbf{m}_2^T\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}^{-1}\mathbf{m}_2. \tag{41}$$

From (12), it follows that

$$\mathbf{m}_1 = -\alpha_2\begin{bmatrix}\mathbf{0}\\ \mathbf{m}\end{bmatrix} = \begin{bmatrix}(M+1)\times 1\\ (M+1)\times 1\end{bmatrix} \tag{42}$$

$$\mathbf{m}_2 = \alpha_1\begin{bmatrix}\mathbf{m}\\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}(M+1)\times 1\\ (M+1)\times 1\end{bmatrix} \tag{43}$$

where

$$\mathbf{m} = \left[\Im\left\{\mathbf{s}_1^H\mathbf{T}^0\mathbf{s}_2\right\}, \ldots, \Im\left\{\mathbf{s}_1^H\mathbf{T}^M\mathbf{s}_2\right\}\right]^T. \tag{44}$$

Now partition the matrix $\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}$ as follows:

$$\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} = \begin{bmatrix} \mathbf{J}_{\boldsymbol{\omega}_1\boldsymbol{\omega}_1} & \mathbf{J}_{\boldsymbol{\omega}_1\boldsymbol{\omega}_2} \\ \mathbf{J}_{\boldsymbol{\omega}_2\boldsymbol{\omega}_1} & \mathbf{J}_{\boldsymbol{\omega}_2\boldsymbol{\omega}_2} \end{bmatrix}. \tag{45}$$

From (11), it is straightforward to show that

$$\mathbf{J}_{\boldsymbol{\omega}_1\boldsymbol{\omega}_1} = \alpha_1^2\Re\left\{\begin{bmatrix} \mathbf{s}_1^H\mathbf{T}^{2\times 0}\mathbf{s}_1 & \ldots & \mathbf{s}_1^H\mathbf{T}^M\mathbf{s}_1 \\ \vdots & \ddots & \vdots \\ \mathbf{s}_1^H\mathbf{T}^M\mathbf{s}_1 & \ldots & \mathbf{s}_1^H\mathbf{T}^{2\times M}\mathbf{s}_1 \end{bmatrix}\right\} \tag{46}$$

$$\mathbf{J}_{\boldsymbol{\omega}_2\boldsymbol{\omega}_2} = \alpha_2^2\Re\left\{\begin{bmatrix} \mathbf{s}_2^H\mathbf{T}^{2\times 0}\mathbf{s}_2 & \ldots & \mathbf{s}_2^H\mathbf{T}^M\mathbf{s}_2 \\ \vdots & \ddots & \vdots \\ \mathbf{s}_2^H\mathbf{T}^M\mathbf{s}_2 & \ldots & \mathbf{s}_2^H\mathbf{T}^{2\times M}\mathbf{s}_2 \end{bmatrix}\right\} \tag{47}$$

$$\mathbf{J}_{\boldsymbol{\omega}_1\boldsymbol{\omega}_2} = \mathbf{J}_{\boldsymbol{\omega}_2\boldsymbol{\omega}_1} = \alpha_1\alpha_2\Re\left\{\begin{bmatrix} \mathbf{s}_1^H\mathbf{T}^{2\times 0}\mathbf{s}_2 & \ldots & \mathbf{s}_1^H\mathbf{T}^M\mathbf{s}_2 \\ \vdots & \ddots & \vdots \\ \mathbf{s}_1^H\mathbf{T}^M\mathbf{s}_2 & \ldots & \mathbf{s}_1^H\mathbf{T}^{2\times M}\mathbf{s}_2 \end{bmatrix}\right\}. \tag{48}$$

Continuing, it is easy to show that

$$\Re\{\mathbf{s}_1^H\mathbf{T}^m\mathbf{s}_1\} = \Re\{\mathbf{s}_2^H\mathbf{T}^m\mathbf{s}_2\} \quad \forall m = 0, 1, \ldots, 2M \tag{49}$$

$$\Re\{\mathbf{s}_1^H\mathbf{T}^m\mathbf{s}_2\} = \Re\{\mathbf{s}_2^H\mathbf{T}^m\mathbf{s}_1\} \quad \forall m = 0, 1, \ldots, 2M. \tag{50}$$

Hence, it follows from (46), (47), and (49) that

$$\frac{1}{\alpha_1^2}\mathbf{J}_{\boldsymbol{\omega}_1\boldsymbol{\omega}_1} = \frac{1}{\alpha_2^2}\mathbf{J}_{\boldsymbol{\omega}_2\boldsymbol{\omega}_2}. \tag{51}$$

But from (42), (43), and (45), and the fact that both $\mathbf{J}_{\boldsymbol{\omega}_1\boldsymbol{\omega}_1}$ and $\mathbf{J}_{\boldsymbol{\omega}_2\boldsymbol{\omega}_2}$ are positive definite, the equality (41) to be proved is equivalent to

$$\alpha_1^2\mathbf{m}^T\left(\mathbf{J}_{\omega_1\omega_1} - \mathbf{J}_{\omega_1\omega_2}\mathbf{J}_{\omega_2\omega_2}^{-1}\mathbf{J}_{\omega_1\omega_2}\right)^{-1}\mathbf{m} = \alpha_2^2\mathbf{m}^T\left(\mathbf{J}_{\omega_2\omega_2} - \mathbf{J}_{\omega_1\omega_2}\mathbf{J}_{\omega_1\omega_1}^{-1}\mathbf{J}_{\omega_1\omega_2}\right)^{-1}\mathbf{m}. \tag{52}$$

The proof of the first part of the corollary follows directly from (51) and (52).

It remains to prove the results for the polynomial phase parameters because the results still hold for the amplitudes from the first part. We will show that when the power is the same, i.e., $\alpha_1 = \alpha_2 = \alpha$, the roles of $\mathbf{s}_1$ and $\mathbf{s}_2$ are equivalent as far as the matrix $\mathbf{C}_{\boldsymbol{\omega}\boldsymbol{\omega}}$ is concerned. To do so, we will make use of the partitioned matrix inversion theorem to obtain

$$\mathbf{C}_{\boldsymbol{\omega}\boldsymbol{\omega}} = \left(\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} - \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\alpha}}\mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}^{-1}\mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\omega}}\right)^{-1}. \tag{53}$$

Using (49) and (50), it can be readily shown from (39) and (46)–(48) that when $\alpha_1 = \alpha_2 = \alpha$, $\mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}$, $\mathbf{J}_{\omega_1\omega_1}$, $\mathbf{J}_{\omega_1\omega_2}$, $\mathbf{J}_{\omega_2\omega_2}$,

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

and hence $\mathbf{J}_{\boldsymbol{\omega\omega}}$ are unchanged when swapping $\mathbf{s}_1$ with $\mathbf{s}_2$. Due to the fact that

$$\Im\left\{\mathbf{s}_1^H\mathbf{T}^m\mathbf{s}_2\right\} = -\Im\left\{\mathbf{s}_2^H\mathbf{T}^m\mathbf{s}_1\right\}, \quad \forall m = 0, 1, \ldots, M \tag{54}$$

$\mathbf{J}_{\boldsymbol{\omega\alpha}}$ and $\mathbf{J}_{\boldsymbol{\alpha\omega}}$ both change sign when swapping $\mathbf{s}_1$ with $\mathbf{s}_2$. However, the product $\mathbf{J}_{\boldsymbol{\omega\alpha}}\mathbf{J}_{\boldsymbol{\alpha\alpha}}^{-1}\mathbf{J}_{\boldsymbol{\alpha\omega}}$ is unchanged. Therefore, it follows from (52) that $\mathbf{C}_{\boldsymbol{\omega\omega}}$ is unchanged when swapping $\mathbf{s}_1$ with $\mathbf{s}_2$. This implies that the Cramér–Rao bounds of the polynomial phase parameters of the two components are the same on a pair-wise basis. This concludes the proof. ■

APPENDIX B
DERIVATION OF (34)

Suppose that the location of the new peak is at $\omega_t = \omega_0 + \delta\omega$ and that $\delta\omega$ is sufficiently small so that we can approximate $\exp\{j\omega_t\} \approx \exp\{j\omega_0\}(1 + j\delta\omega)$. By definition, one can show that ω_t is the solution of

$$\frac{\partial g(\omega)}{\partial \omega} g(\omega) = -\frac{\partial h(\omega)}{\partial \omega} g(\omega) \tag{55}$$

where $g(\omega) = \Re\{Y_2(\omega)\}$ and $h(\omega) = \Im\{Y_2(\omega)\}$. By noting that $g(\omega)$ and $h(\omega)$ are real-valued functions of ω, it can be deduced that

$$g'(\omega) = \Re\left\{\frac{\partial Y_2(\omega)}{\partial \omega}\right\}, \quad h'(\omega) = \Im\left\{\frac{\partial Y_2(\omega)}{\partial \omega}\right\}. \tag{56}$$

It is easily shown that

$$Y_2(\omega_t) \approx \sum_{n=0}^{N-1} y_2(n)\exp\{-j\omega_0\}(1 - jn\delta\omega) \tag{57}$$

$$Y_2'(\omega_t) \approx \sum_{n=0}^{N-1} -jny_2(n)\exp\{-j\omega_0\}(1 - jn\delta\omega). \tag{58}$$

Denote $p(n) = \Re\{y_2(n)\exp\{-j\omega_0\}\}$, $q(n) = \Im\{y_2(n)\exp\{-j\omega_0\}\}$. One can deduce from the above equations that

$$g(\omega_t) = \sum_{n=0}^{N-1} p(n) + nq(n)\delta\omega \tag{59}$$

$$g'(\omega_t) = \sum_{n=0}^{N-1} nq(n) - n^2p(n)\delta\omega \tag{60}$$

$$h(\omega_t) = \sum_{n=0}^{N-1} q(n) - np(n)\delta\omega \tag{61}$$

$$h'(\omega_t) = \sum_{n=0}^{N-1} -np(n) - n^2q(n)\delta\omega \tag{62}$$

from which we arrive at (34).

REFERENCES

[1] T. J. Abatzoglou, "Fast maximum likelihood joint estimation of frequency and frequency rate," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 708–715, 1986.

[2] J. Ängeby, "Aliasing of polynomial-phase signal parameters," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1488–1491, 2000.

[3] ——, "Estimating signal parameters using the nonlinear instantaneous least squares approach," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2721–2732, 2000.

[4] S. Barbarossa, "Analysis of multicomponent LFM signals by a combined Wigner-Hough transform," *IEEE Trans. Signal Process.*, vol. 43, pp. 1511–1515, Jun. 1995.

[5] S. Barbarossa and V. Petrone, "Analysis of polynomial-phase signals by the integrated generalised ambiguity function," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 316–327, 1997.

[6] S. Barbarossa, A. Scaglione, and G. B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial phase signal modeling," *IEEE Trans. Signal Process.*, vol. 46, pp. 691–708, Mar. 1998.

[7] P. M. Djuric and S. M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2118–2126, 1990.

[8] J. M. Francos and B. J. Friedlander, "Bounds for estimation of multicomponent signals with random amplitude and deterministic phase," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1161–1172, 1995.

[9] B. Friedlander and J. M. Francos, "Estimation of amplitudes and phase parameters of multicomponent signals," *IEEE Trans. Signal Process.*, vol. 43, no. 4, pp. 917–926, 1995.

[10] M. Ghogho, A. K. Nandi, and A. Swami, "Cramér-Rao bounds and maximum likelihood estimation for random amplitude phased-modulated signals," *IEEE Trans. Signal Process.*, vol. 47, pp. 2905–2916, Nov. 1999.

[11] S. Golden and B. Friedlander, "A modification to the discrete polynomial phase transform," *IEEE Trans. Signal Process.*, vol. 46, pp. 1452–1455, May 1998.

[12] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM J. Optim.*, vol. 9, no. 1, pp. 112–147, 1998.

[13] J. A. Legg and D. A. Gray, "Performance bounds for polynomial phase parameter estimation with nonuniform and random sampling schemes," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 331–337, 2000.

[14] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1984.

[15] M. R. Morelande and A. M. Zoubir, "Model selection of random amplitude polynomial phase signals," *IEEE Trans. Signal Process.*, vol. 50, pp. 578–589, Mar. 2002.

[16] ——, "On the performance of cyclic moments-based parameter estimators of amplitude modulated polynomial phase signals," *IEEE Trans. Signal Process.*, vol. 50, pp. 590–606, Mar. 2002.

[17] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, pp. 308–313, 1965.

[18] P. O'Shea, "A fast algorithm for estimating the parameters of a quadratic FM signal," *IEEE Trans. Signal Process.*, vol. 52, pp. 385–393, Feb. 2004.

[19] S. Peleg and B. Friedlander, "The discrete polynomial-phase transform," *IEEE Trans. Signal Process.*, vol. 43, pp. 1901–1914, Aug. 1995.

[20] ——, "Multicomponent signal analysis using the polynomial phase transform," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, pp. 378–387, Jan. 1996.

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

[22] ——, "The Cramér-Rao lower bound for signals with constant amplitude and polynomial phase," *IEEE Trans. Signal Process.*, vol. 39, pp. 749–752, Mar. 1991.

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

[24] B. Porat and B. Friedlander, "Asymptotic statistical analysis of the high-order ambiguity function for parameter estimation of polynomial-phase signals," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 995–1001, 1996.

[25] W. H. Press, S. A. Teukolsky, W. T. Vettterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[26] A. Scaglione and S. Barbarossa, "On the spectral properties of polynomial-phase signals," *IEEE Signal Process. Lett.*, vol. 5, no. 9, pp. 237–240, 1998.

[27] ——, "Statistical analysis of the product of high-order ambiguity function," *IEEE Trans. Inf. Theory*, vol. 45, pp. 343–356, Jan. 1999.

[28] L. L. Scharf, *Statistical Signal Processing: Detection Estimation and Time Series Analysis*. New York: Addison-Wesley, 1991.

[29] S. Shamsunder, G. B. Giannakis, and B. Friedlander, "Estimating random amplitude polynomial phase signals: a cyclosationary approach," *IEEE Trans. Signal Process.*, vol. 43, pp. 492–505, Feb. 1995.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.

[30] C. F. Wu, "Asymptotic theory of nonlinear least squares estimation," *Ann. Statist.*, vol. 9, no. 3, pp. 501–513, May 1981.
[31] G. T. Zhou, G. B. Giannakis, and A. Swami, "On polynomial phase signals with time-varying amplitudes," *IEEE Trans. Signal Process.*, vol. 44, pp. 848–861, Apr. 1996.
[32] A. M. Zoubir and D. R. Iskander, "Bootstrap modeling of a class of nonstationary signals," *IEEE Trans. Signal Process.*, vol. 48, pp. 399–408, Feb. 2000.

Duc Son Pham (S'02–M'05) received the B.Eng. degree (with first-class honors) from Monash University, Australia, in 2001 and the Ph.D. degree in electrical engineering from Curtin University of Technology, Perth, Australia, in 2005.

He is now with Curtin University of Technology as a Postdoctoral Research Fellow. His current research interests fall in the area of statistical signal processing with application to communications.

Abdelhak M. Zoubir (SM'97) received the Dipl.-Ing degree from Fachhochschule Niederrhein, Germany, in 1983 and the Dipl.-Ing. and Dr.-Ing. degrees from Ruhr University, Bochum, Germany, in 1987 and 1992, respectively, all in electrical engineering.

Early placement in industry (Klöckner-Moeller and Siempelkamp AG) was followed by an Associate Lectureship in the Division for Signal Theory, Ruhr University. In 1992, he joined Queensland University of Technology, Australia, where he was a Lecturer, Senior Lecturer, and then Associate Professor in the School of Electrical and Electronic Systems Engineering. In 1999, he became a Professor of telecommunications at Curtin University of Technology, where he was Head of the School of Electrical and Computer Engineering from November 2001 until February 2003. In 2003, he became a Professor of signal processing at Darmstadt University of Technology. His general research interest lies in statistical methods for signal processing with applications in communications, sonar, radar, biomedical engineering, and vibration analysis. His current research interest lies in robust estimation and in bootstrap techniques for spectrum estimation and the modelling of nonstationary and non-Gaussian signals. He is currently Associate Editor of *Signal Processing* and the *Journal of Applied Signal Processing*.

Prof. Zoubir is a member of the Institute of Mathematical Statistics. He was General Cochairman of the Third IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Darmstadt, in December 2003, Technical Chairman of the 11th IEEE Workshop on Statistical Signal Processing, Singapore, in August 2001, and Deputy Technical Chairman (Special Sessions/Tutorials) of ICASSP-94, Adelaide, Australia. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 until 2005. He is a Member of the IEEE SPS Technical Committees on Signal Processing Theory and Methods and Signal Processing Education.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on October 14, 2009 at 23:28 from IEEE Xplore. Restrictions apply.