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].

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 autoterm 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 autoterm 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
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-PPSSs. 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 $\sigma^2$

$$y(t) = \sum_{k=1}^{K} s_k(t) + \nu(t)$$

$$= \sum_{k=1}^{K} \alpha_k \exp \left\{ j \sum_{m=0}^{M} \omega_k^m t^m \right\} + \nu(t), \quad t \in \mathbb{R}. \quad (1)$$

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 $\alpha_k$ and the polynomial phase coefficients $\omega_k^m$ 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 $\omega_k = [\omega_k^0, \ldots, \omega_k^M]^T$, $t = [t_0, \ldots, t_{N-1}]^T$, $y = [y(t_0), \ldots, y(t_{N-1})]^T$, $\alpha = [\alpha_1, \ldots, \alpha_K]^T$, $\nu = [\nu(t_0), \ldots, \nu(t_{N-1})]^T$, $s_k(t, \omega_k) = \exp \left\{ j \sum_{m=0}^{M} \omega_k^m t^m \right\}$, $s_k(t_0, \omega_k), \ldots, s_k(t_{N-1}, \omega_k)]^T = s_k(\omega_k)^T$, $\omega = (\omega_1, \ldots, \omega_K)$, and $S(\omega) = [s_1(\omega_1), \ldots, s_K(\omega_K)]$. We can rewrite (1) as

$$y = [s_1(t, \omega_1), \ldots, s_K(t, \omega_K)] \alpha + \nu$$

$$= (S(\omega)\alpha + \nu). \quad (2)$$

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

$$p(y; \theta) = \frac{1}{(\sigma^2)^N} \exp \left\{ -\frac{(y - S(\omega)\alpha)^H (y - S(\omega)\alpha)}{\sigma^2} \right\}. \quad (3)$$

The log-likelihood function is

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

The $(i, j)^{th}$ element of Fisher’s information matrix is given by

$$J_{ij} = -\mathbb{E} \left\{ \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta_i \partial \theta_j} \right\}. \quad (4)$$

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

$$J = \begin{bmatrix} J_{\alpha\alpha} & J_{\alpha\omega} & 0 \\ J_{\omega\alpha} & J_{\omega\omega} & 0 \\ 0 & 0 & J_{\sigma^2, \sigma^2} \end{bmatrix}. \quad (5)$$

Recall that

$$J_{ij} = \text{tr} \left\{ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \theta_i} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \theta_j} \right\} + 2 \text{tr} \left\{ \frac{\partial \mathbf{H}}{\partial \theta_i} \mathbf{R}^{-1} \frac{\partial \mathbf{H}}{\partial \theta_j} \right\} \quad (6)$$

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

$$\mu = S(\omega)\alpha$$

$$= \sum_{k=1}^{K} s_k(\omega_k)\alpha_k. \quad (7)$$

It can be easily shown that

$$\frac{\partial \mu}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k} \sum_{k=1}^{K} s_k(\omega_k)\alpha_k$$

$$= s_k(\omega_k). \quad (8)$$

Therefore

$$J_{\alpha\alpha} = \frac{2}{\sigma^2} \text{tr} \left\{ S(\omega)^H S(\omega) \right\}. \quad (9)$$

Now, for the block $J_{\omega\omega}$, we note that

$$\frac{\partial \mu}{\partial \omega_k} = \frac{\partial}{\partial \omega_k} \sum_{k=1}^{K} s_k(t, \omega_k)\alpha_k$$

$$= \sum_{k=1}^{K} \alpha_k \exp \left\{ j \sum_{m=0}^{M} \omega_k^m t^m \right\}$$

$$= j \omega_k^T \mathbf{T} s_k(\omega_k) \quad (10)$$

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

$$J_{\omega\omega} \left( \omega_k^m, \omega_k^{m'} \right) = \frac{2}{\sigma^2} \text{tr} \left\{ \alpha_k s_k(\omega_k)^H \mathbf{T}^{m+m'} s_k(\omega_k) \right\}. \quad (11)$$

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

$$J_{\omega\omega} \left( \alpha_k, \omega_k \right) = \frac{2}{\sigma^2} \text{tr} \left\{ \alpha_k s_k(\omega_k)^H \mathbf{T} s_k(\omega_k) \right\}. \quad (12)$$

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

$$J_{\sigma^2, \sigma^2} = N/\sigma^4. \quad (13)$$

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

The parameter $t$ is omitted for notational simplification.
There are some observations on the Cramér–Rao bound for mc-PPSSs.

- Compared with the results for monocomponent PPSs [22], [23], the Cramér–Rao bound for mc-PPSSs 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-PPSSs 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 \( \alpha_1 \) and \( \alpha_2 \), i.e., the diagonal elements of \( \mathbf{C} \) that correspond to \( \alpha_1 \) and \( \alpha_2 \) are the same

\[
\text{CRB}\{\alpha_1\} = \text{CRB}\{\alpha_2\} \quad (14)
\]

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.,

\[
\text{CRB}\{\omega_m^1\} = \text{CRB}\{\omega_m^2\}, \quad m = 0, 1, \ldots, M. \quad (15)
\]

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-PPSSS

#### 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:

\[
y_1(t) = y(t)
\]

\[
y_2(t, t_1) = y_1(t + t_1)y_1(t - t_1) \quad (16)
\]

\[
y_3(t, t_1, t_2) = y_1(t + t_1 + t_2)y_1(t + t_1 - t_2)y_1(t - t_1 + t_2)y_1(t - t_1 - t_2) \quad (17)
\]

\[
\vdots
\]

\[
y_M(t; \tau_{M-1}) = y_{M-1}(t + \tau_{M-1}; \tau_{M-1}) \times y_{M-1}(t - \tau_{M-1}; \tau_{M-1})^* \quad (18)
\]

where \( \tau_i = [\tau_1, \tau_2, \ldots, \tau_{M-1}] \). 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

\[
Y_M(f; \tau_{M-1}) = \sum_{l=0}^{\frac{M-1}{2}} y_M(t; \tau_{M-1})e^{-j2\pi ft}. \quad (20)
\]

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 \( M \)-th order HIM contains only a single sinusoid whose frequency is linearly proportional to the highest phase coefficient \( \omega_M \). Hence, one can perform peak picking over the discrete Fourier transform (DFT) of the HIM and estimate \( \omega_M \). 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 \( M \)-th order PPS, one can show by direct evaluation that the ml-HIM contains \( K \) sinusoids that correspond to \( K \) autoterm, each having a frequency proportional to the highest order phase coefficient \( \omega_M^K \). However, the ml-HIM also contains a large number of cross-terms which are also \( M \)-th 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-PPSSs. Except for second-order mc-PPSSs 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 \( m_1 + \ldots + m_{M-1} = M \), 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 autoterm has the peculiar property that they are single sinusoids and their frequencies are proportional to \( \omega_M^K \) and the product of lags. Hence, by using different sets of lags and after proper scaling and multiplications of scaled ml-HAFs, the autoterm are always enhanced more significantly than the cross-terms, which strengthen the discriminant ability. Mathematically, suppose that we have \( L \) sets of lags

\[
T_{M-1}^{(l)} = \left[ \tau_{M-1}^{(1)}, \tau_{M-1}^{(2)}, \ldots, \tau_{M-1}^{(L)} \right]^T, \quad (21)
\]

The PHAF is defined as follows:

\[
Y_M^{(l)}(f; T_{M-1}^{(l)}) = \prod_{l=1}^{L} Y_M\left(\beta^{(l)} f; T_{M-1}^{(l)}\right) \quad (22)
\]

where the scale factor is

\[
\beta^{(l)} = P_M \left( T_{M-1}^{(l)}, T_{M-1}^{(l)} \right) = \frac{\left( \prod_{l=1}^{M-1} \tau_{M-1}^{(l)} \right)}{\left( \prod_{l=1}^{M-1} \tau_{M-1}^{(l)} \right)^{1/M}}. \quad (23)
\]

The introduction of the scale factor is to align the autoterm from different sets of lags. It is noted that the mechanism of the PHAF is analogous to integration along the curve of \( Y_M^{(l)}(f; T_{M-1}^{(l)}) \) in the ambiguity domain whose two axes are
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 \( t_n = n \)

\[
y(n) = \alpha_1 \exp \left\{ j \left( \omega_0^1 n + \omega_1^2 n^2 \right) \right\} + \alpha_2 \exp \left\{ j \left( \omega_0^2 n^2 + \omega_1 n^2 \right) \right\}.
\]

(24)

Suppose that the first component is of interest. We shall describe a perturbation analysis on the estimate of \( \omega_2^1 \) 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 a 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

\[
y_2(n) = y(n + k) y(n - k)^*.
\]

(25)

The HAF is

\[
Y_2(\omega) = \sum_{n=0}^{N-1} y_2(n) \exp \{-j\omega n\}.
\]

(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 \( \omega_0 \). 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 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 \( \alpha \) and \( \omega \) in (2)

\[
(\hat{\alpha}, \hat{\omega}) = \arg \min_{\alpha, \omega} (y - S(\omega) \alpha)^H (y - S(\omega) \alpha).
\]

(27)

To reduce the number of unknowns in (27), we further split the polynomial phase parameters into \( \omega_k = (\omega_0^k, \theta_M^k) \), \( k = 1, \ldots, K \), and introduce \( \theta = (\theta_1, \ldots, \theta_K) \). Essentially, \( \theta_k \) consists of \( M \) polynomial phase parameters with orders ranging from one to \( M \). Introduce \( \eta = [\omega_0^1, \omega_0^2, \ldots, \omega_0^K]^T \) and decompose

\[
S(\omega) = H(\theta) c(\eta, \alpha)
\]

(28)

where

\[
c(\eta, \alpha) = \begin{bmatrix} \alpha_1 \exp \{ j \omega_0^1 \} & \ldots & \alpha_K \exp \{ j \omega_0^K \} \end{bmatrix}^T \in \mathbb{C}^K
\]

(29)

\[
H(\theta) = [h_1(\theta_1), \ldots, h_K(\theta_K)] \in \mathbb{C}^{N \times K}
\]

(30)

and \( h_k(\theta_k) \) is a column vector whose \( m \)th element is \( \exp \{ j \sum_{l=1}^{m} \omega_0^k l \} \). Note that, when the noise is complex circular white Gaussian and if the estimate of \( \theta \), which we denote by \( \hat{\theta} \), is available, then \( \hat{\eta} \) and \( \hat{\alpha} \) are estimated from

\[
\hat{\epsilon} = H(\hat{\theta})^\dagger y
\]

(31)

\[
\hat{\alpha} = |\hat{\epsilon}|
\]

(32)

\[
\hat{\eta} = \zeta \{ \hat{\epsilon} \}
\]

(33)

where \( H(\hat{\theta})^\dagger = (H(\hat{\theta})^H H(\hat{\theta}))^{-1} H(\hat{\theta})^H \) denotes the pseudoinverse of \( H(\hat{\theta}) \). 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:

\[
\hat{\theta} = \arg \min_{\theta} f(\theta)
\]

(35)

\[
\hat{\theta} = \arg \min_{\theta} \{ -y^H P \hat{H} \hat{H}^H H(\hat{\theta}) y \}
\]

where \( P \hat{H} \hat{H}^H = H(\hat{\theta}) (H(\hat{\theta})^H H(\hat{\theta}))^{-1} H(\hat{\theta})^H \) denotes the projection matrix onto the subspace of \( H(\hat{\theta}) \). Asymptotic theory of NLS estimation can be found in, for example, [30]. Note that 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 NLS 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 NLS 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),

\[
\delta \omega \approx \frac{\sum n p(n) (\sum q(n)) - (\sum n q(n)) (\sum p(n))}{(\sum n q(n))^2 + (\sum n^2 p(n))^2 - (\sum n^2 q(n)) (\sum n^2 p(n)) (\sum n^2 q(n)) (\sum q(n))}
\]

(34)
TABLE I
EXAMPLE OF A SECOND ORDER mc-PPS

<table>
<thead>
<tr>
<th></th>
<th>True values</th>
<th>PHAF estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Component 1</td>
<td>Component 2</td>
</tr>
<tr>
<td></td>
<td>Component 1</td>
<td>Component 2</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1</td>
<td>0.99821699938630</td>
</tr>
<tr>
<td>$\omega_0$</td>
<td>0</td>
<td>-0.07853902562474</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>2.19911485751286</td>
<td></td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>0.00490873852123</td>
<td></td>
</tr>
</tbody>
</table>

1) Initialization: The PHAF method can be used to initialize the estimates with some modifications. After estimating the highest order phase coefficients $\omega_{k1}^f$, $k = 1, \ldots, K$, 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 $f(\theta), \theta \in \mathbb{R}^n$ is to be minimized with a starting point $\theta_0$. The $k$th iteration of the NM algorithm for solving this problem is summarized in Table II, where the parameters are reflection $r = 1$, 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.

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 $\alpha_1 = \alpha_2 = 1$, $(\omega_{11}^f, \omega_{12}^f, \omega_{22}^f) = (0, 20\pi, -0.22\pi/N)$, $(\omega_{11}^0, \omega_{12}^0, \omega_{22}^0) = (0, 0.80\pi, -0.31\pi/N)$, 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 PHAF 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 $\omega_{22}^f$ 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.
In the simulation settings, we selected maximum computer accuracy (i.e., the tolerance is $2.204 \times 10^{-16}$). 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 when $\omega_0$ 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 $\geq 5$ dB and investigate the MSE performance.

- In the simulation settings, we selected maximum computer accuracy (i.e., the tolerance is $2.204 \times 10^{-16}$). 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 $\geq 5$ dB and investigate the MSE performance.
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 PHAF method suffers from a considerable finite sample error 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 $\alpha_1 = \alpha_2 = 1$, $(\omega_1^1, \omega_2^1, \omega_3^1, \omega_4^1) = (0, 20\pi, -0.10\pi/N, 0.30\pi/N^2)$, $(\omega_1^2, \omega_2^2, \omega_3^2, \omega_4^2) = (0.07\pi, 0.40\pi/N, -0.30\pi/N^2)$. $N = 128$. 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 $\omega_3^1$). 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 $N$. 

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-PPSSs, 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 \(s_1\) for \(s_1(\omega_1)\) and \(s_2\) for \(s_2(\omega_2)\).
- Ignore \(J_{\omega_2\omega_2}\), as it does not affect subsequent results. Hence, Fisher’s information matrix now becomes

\[
J = \begin{bmatrix}
J_{\omega\omega} & J_{\omega\alpha} \\
J_{\alpha\omega} & J_{\alpha\alpha}
\end{bmatrix}.
\]  

(36)

The Cramér–Rao bound is therefore

\[
C = J^{-1} = \begin{bmatrix}
C_{\omega\omega} & C_{\omega\alpha} \\
C_{\alpha\omega} & C_{\alpha\alpha}
\end{bmatrix}
\]  

(37)

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

\[
C_{\omega\alpha} = (J_{\omega\alpha} - J_{\omega\omega}^{-1}J_{\omega\omega}J_{\omega\alpha})^{-1}.
\]

(38)

It follows from (9) that

\[
J_{\omega\alpha} = \begin{bmatrix}
N & \Re \{s_1^H s_2\} \\
\Re \{s_1^H s_2\} & N
\end{bmatrix}.
\]

(39)

Denote

\[
J_{\omega\omega} = \begin{bmatrix} m_1 & m_2 \end{bmatrix}.
\]

(40)

Using the fact that \(J_{\omega\omega}\) 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 \(C_{\omega\alpha}\) to be the same is

\[
m_1^T J_{\omega\alpha}^{-1} m_1 = m_2^T J_{\omega\alpha}^{-1} m_2.
\]

(41)

From (12), it follows that

\[
m_1 = -\alpha_2 \begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix}(M+1) \times 1\end{bmatrix}
\]

(42)

\[
m_2 = \alpha_1 \begin{bmatrix} m \end{bmatrix} = \begin{bmatrix}(M+1) \times 1\end{bmatrix}
\]

(43)

where

\[
m = \begin{bmatrix} 3 \{s_1^H T_0 s_2\} , & \ldots & 3 \{s_1^H T_M s_2\} \end{bmatrix}^T.
\]

(44)

Now partition the matrix \(J_{\omega\omega}\) as follows:

\[
J_{\omega\omega} = \begin{bmatrix}
J_{\omega_1\omega_1} & J_{\omega_1\omega_2} \\
J_{\omega_2\omega_1} & J_{\omega_2\omega_2}
\end{bmatrix}.
\]

(45)

From (11), it is straightforward to show that

\[
J_{\omega_1\omega_1} = \begin{bmatrix}
\Re \{s_1^H T_0 s_1\} & \ldots & \Re \{s_1^H T_M s_1\} \\
\ldots & \ldots & \ldots
\end{bmatrix}
\]

\[
J_{\omega_2\omega_2} = \begin{bmatrix}
\Re \{s_2^H T_0 s_2\} & \ldots & \Re \{s_2^H T_M s_2\} \\
\ldots & \ldots & \ldots
\end{bmatrix}
\]

\[
J_{\omega_1\omega_2} = \begin{bmatrix}
\Re \{s_1^H T_0 s_2\} & \ldots & \Re \{s_1^H T_M s_2\} \\
\ldots & \ldots & \ldots
\end{bmatrix}
\]

\[
J_{\omega_2\omega_1} = \begin{bmatrix}
\Re \{s_2^H T_0 s_1\} & \ldots & \Re \{s_2^H T_M s_1\} \\
\ldots & \ldots & \ldots
\end{bmatrix}
\]

(46)

(47)

(48)

Continuing, it is easy to show that

\[
\Re \{s_1^H T_m s_1\} = \Re \{s_2^H T_m s_2\} \quad \forall m = 0, 1, \ldots, 2M
\]

(49)

\[
\Re \{s_1^H T_m s_2\} = \Re \{s_2^H T_m s_1\} \quad \forall m = 0, 1, \ldots, 2M.
\]

(50)

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

\[
\frac{1}{\alpha_1^2} J_{\omega_1\omega_1} = \frac{1}{\alpha_2^2} J_{\omega_2\omega_2}.
\]

(51)

But from (42), (43), and (45), and the fact that both \(J_{\omega_1\omega_1}\) and \(J_{\omega_2\omega_2}\) are positive definite, the equality (41) to be proved is equivalent to

\[
\alpha_2^2 \begin{bmatrix} m^T (J_{\omega_1\omega_1} - J_{\omega_1\omega_2} J_{\omega_2\omega_2}^{-1} J_{\omega_1\omega_1})^{-1} m \\
\end{bmatrix} = \alpha_1^2 \begin{bmatrix} m^T (J_{\omega_2\omega_2} - J_{\omega_2\omega_1} J_{\omega_1\omega_2}^{-1} J_{\omega_2\omega_1})^{-1} m
\]

(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 \(s_1\) and \(s_2\) are equivalent as far as the matrix \(C_{\omega\omega}\) is concerned. To do so, we make use of the partitioned matrix inversion theorem to obtain

\[
C_{\omega\omega} = (J_{\omega\omega} - J_{\omega\omega} J_{\omega\omega}^{-1} J_{\omega\omega})^{-1}.
\]

(53)

Using (49) and (50), it can be readily shown from (39) and (46)–(48) that when \(\alpha_1 = \alpha_2 = \alpha\), \(J_{\omega_1\omega_1} = J_{\omega_1\omega_2}, J_{\omega_2\omega_1} = J_{\omega_2\omega_2}\),
and hence $J_{uo\omega}$ are unchanged when swapping $s_1$ with $s_2$. Due to the fact that
\[ \Im \{ s_1^H T^m s_2 \} = -\Im \{ s_2^H T^m s_1 \}, \quad \forall m = 0, 1, \ldots, M \]
(54)
$J_{uo\omega}$ and $J_{ao\omega}$ both change sign when swapping $s_1$ with $s_2$. However, the product $J_{uo\omega} J_{ao\omega}$ is unchanged. Therefore, it follows from (52) that $C_{uo\omega}$ is unchanged when swapping $s_1$ with $s_2$. 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 $\omega_k = \omega_0 + \delta \omega$ and that $\delta \omega$ is sufficiently small so that we can approximate $\exp\{j\omega_k t\} \approx \exp\{j\omega_0 t\}(1 + j\delta \omega t)$. By definition, one can show that $\omega_k$ is the solution of
\[ \frac{\partial g(\omega)}{\partial \omega} g(\omega) = -\frac{\partial h(\omega)}{\partial \omega} h(\omega) \]
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 $\omega$, it can be deduced that
\[ g'(\omega) = \Re \left\{ \frac{\partial Y_2(\omega)}{\partial \omega} \right\}, \quad h'(\omega) = \Im \left\{ \frac{\partial Y_2(\omega)}{\partial \omega} \right\} \]
(56)
It is easily shown that
\[ Y_2(\omega_k) \approx \sum_{n=0}^{N-1} y(n) \exp\{-j\omega_k t\}(1 - jn \delta \omega) \]
(57)
\[ Y_2'(\omega_k) \approx \sum_{n=0}^{N-1} -jny(n) \exp\{-j\omega_k t\}(1 - jn \delta \omega) \]
(58)
Denote $p(n) = \Re \{ y(n) \exp\{-j\omega_0 t\} \}$, $q(n) = \Im \{ y(n) \exp\{-j\omega_0 t\} \}$. One can deduce from the above equations that
\[ g(\omega_k) = \sum_{n=0}^{N-1} p(n) + nq(n) \delta \omega \]
(59)
\[ g'(\omega_k) = \sum_{n=0}^{N-1} nq(n) - n^2 p(n) \delta \omega \]
(60)
\[ h(\omega_k) = \sum_{n=0}^{N-1} q(n) - np(n) \delta \omega \]
(61)
\[ h'(\omega_k) = \sum_{n=0}^{N-1} -np(n) - n^2 q(n) \delta \omega \]
(62)
from which we arrive at (34).

**REFERENCES**


Duc Son Pham (S’02–M’05) received the B.Eng.
degree (with first-class honors) from Monash Uni-
versity, Australia, in 2001 and the Ph.D. degree in
electrical engineering from Curtin University of
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 Asso-
ciate 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.