Progress In Electromagnetics Research C, Vol. 23, 111–122, 2011

MAXIMUM LIKELIHOOD ESTIMATION OF CO-CHANNEL MULTICOMPONENT POLYNOMIAL PHASE SIGNALS USING IMPORTANCE SAMPLING

H. Cheng*, D. Zeng, J. Zhu, and B. Tang

School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Abstract—Unlike some traditional polynomial phase signal (PPS) parameter estimation methods restricted to monocomponent case, this paper focuses on the parameter estimation of multicomponent PPSs mixed in a single channel, which is more sophisticated and always involves the cross-term issue. In this investigation, based on the model of multicomponent PPSs in additional white Gaussian noise, we partition the maximum likelihood estimation into two consecutive steps. The first one involving estimation of polynomial coefficients is intensively studied using importance sampling, while the second one involving the estimation of amplitude and initial phase is trivial. Numerical experiments show satisfactory estimation performance even if the parameters are closely spaced.

1. INTRODUCTION

The polynomial phase signal (PPS) has been flourishing in radar in recent years. Moreover, in fields such as radio communications and sonar technology, the involved signals are always nonstationary and have the property of continuous instantaneous phase. Based on the Weierstrass' theorem, these signals can be well approximated by a PPS with finite order phase, if the instantaneous phase is within a closed interval.

Thus far, most parameter estimation algorithms of PPS have focused on monocomponent case. These algorithms can be reduced into two categories, the estimation from the statistical theory and the one from a transform. Algorithms from the former category include maximum likelihood (ML) estimation [1], nonlinear instantaneous least

Received 20 June 2011, Accepted 28 July 2011, Scheduled 7 August 2011
* Corresponding author: Hao Cheng (haoc08@gmail.com).

squares (NILS) [2], nonlinear least squares (NLS) [3], and others, while those from the latter are mainly based on the principle of the polynomial phase transform (PPT) [4–6].

Nevertheless, both above-mentioned categories have issues when directly applied to estimate the parameters of the multicomponent PPSs mixed in a single channel. Firstly, algorithms from statistical theories, for example, the ML estimation, will suffer more computational burden for its multidimensional grid search than in the case of monocomponent, or need an appropriate guess of the initial value when using an iterative optimizer. Secondly, algorithms from PPT would be affected by the cross-term issue, which arguably exists in multicomponent case. For the parameter estimation of multicomponent PPSs, literature [7] gives some theoretical analysis using the PPT.

Motivated by the assumption that the likelihood function in ML estimation is regarded as a pseudo probability density function (PDF), from which the optimization problem could be solved using drawn samples, this paper develops a ML parameter estimation algorithm of co-channel multicomponent PPSs using importance sampling, which is utilized for generating samples from a complex PDF. It extends the studies of [8] and [9], which used importance sampling to estimate frequencies of multicomponent sinusoids and estimate the initial frequencies and chirp rates of chirp signals, respectively. Moreover, this method is also referred to as mean likelihood estimation in [8].

The remainder of this paper is organized as follows. Section 2 introduces a different expression of the PPS whose polynomial coefficients (PCs) are converted into a common scale for convenience, and partitions ML estimation into two consecutive steps. Section 3 derives the importance sampling estimator of multicomponent PPSs. Section 4 summarizes the steps of the algorithm. Section 5 gives the Cramer-Rao lower bounds (CRLBs) of multicomponent PPSs. Section 6 presents the experimental results of our algorithm and is followed by the conclusion in Section 7.

2. PROBLEM FORMULATION

2.1. Signal Model with Normalized PCs

The signal model for the superimposed PPSs with the discrete time in additional white Gaussian noise is

$$x(n) = \sum_{p=1}^{P} A^{(p)} e^{j\left(\phi_0^{(p)} + 2\pi \sum_{k=1}^{K} c^{(p,k)} \left(\frac{n}{f_s}\right)^k\right)} + w(n) \tag{1}$$

where $n = 0, 1, \ldots, N-1$, with N denoting the number of data samples, $A^{(p)}$ and $\phi_0^{(p)}$ are the amplitude and the initial phase of the pth PPS, respectively, $c^{(p,k)}$ is the kth order PC of the pth PPS, f_s is the sampling rate, and $w(n)$ denotes white Gaussian noise with zero mean and variance σ^2. Unlike expressions which appear in some other literatures, for the convenience of our algorithm, first we normalize each PC $c^{(p,k)}$ to a common scale as opposed to the sampling rate. For a single PPS, assuming ϕ_0 and $c^{(k)}$ denote the initial phase and the kth order PC, respectively, the instantaneous phase is presented as

$$\phi(n) = \phi_0 + 2\pi \sum_{k=1}^{K} \frac{c^{(k)}}{f_s^k} n^k, \tag{2}$$

and the normalized instantaneous frequency is

$$f(n) = \frac{\phi'(n)}{2\pi} = \sum_{k=1}^{K} k \frac{c^{(k)}}{f_s^k} n^{k-1} = \sum_{k=1}^{K} d^{(k)} \left(\frac{n}{N}\right)^{k-1} \tag{3}$$

where the mark $'$ means derivative and $d^{(k)} = kN^{k-1}c^{(k)} / f_s^k$ is defined as the normalized PC which has the same scale as opposed to the sampling rate.

Assuming each $d^{(k)} > 0$, the normalized instantaneous frequency $f(n)$ is within the range from $d^{(1)} (n = 0)$ to $d^{(1)} + d^{(2)} + \ldots + d^{(K)} (n = N-1)$, where N is assumed to be large enough. Specifically, for $K = 2$, the PPS is a chirp signal with a normalized initial frequency $d^{(1)}$ and a normalized bandwidth $d^{(2)}$. Without loss of generality, in the following, f_s is fixed to 1 and the term 'normalized' is omitted.

Thus, using this expression, the signal model could be expressed as

$$x(n) = \sum_{p=1}^{P} A^{(p)} e^{j\left(\phi_0^{(p)} + 2\pi \sum\limits_{k=1}^{K} d^{(p,k)} \frac{n^k}{kN^{k-1}}\right)} + w(n) \tag{4}$$

where $d^{(p,k)} = kN^{k-1}c^{(p,k)} / f_s^k$. The matrix-vector form of (4) is

$$\mathbf{x} = \mathbf{H}(\mathbf{D})\boldsymbol{\theta} + \mathbf{w} \tag{5}$$

where $\mathbf{x} = [x(0), x(1), \ldots, x(N-1)]^T$, $\mathbf{w} = [w(0), w(1), \ldots, w(N-1)]^T$, $\boldsymbol{\theta} = [A^{(1)} \exp(j\phi_0^{(1)}), A^{(2)} \exp(j\phi_0^{(2)}), \ldots, A^{(P)} \exp(j\phi_0^{(P)})]^T$, $\mathbf{D} = [\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \ldots, \mathbf{d}^{(K)}]$ with $\mathbf{d}^{(k)} = [d^{(1,k)}, d^{(2,k)}, \ldots, d^{(P,k)}]^T$ for $k = 1, 2, \ldots, K$, and $\mathbf{H}(\mathbf{D}) = [\mathbf{h}^{(1)}(\mathbf{D}), \mathbf{h}^{(2)}(\mathbf{D}), \ldots, \mathbf{h}^{(P)}(\mathbf{D})]$ with $\mathbf{h}^{(p)}(\mathbf{D}) = [1, \exp(j(2\pi \sum\limits_{k=1}^{K} d^{(p,k)} / kN^{k-1})), \ldots, \exp(j(2\pi \sum\limits_{k=1}^{K} d^{(p,k)} (N-1)^k / kN^{k-1}))]^T$ for $p = 1, 2, \ldots, P$.

2.2. ML Estimation

The PDF of the received data is

$$p\left(\mathbf{x};\mathbf{D},\boldsymbol{\theta}\right)=\frac{1}{\pi^{N}\sigma^{2N}}\exp\left[-\frac{1}{\sigma^{2}}\left(\mathbf{x}-\mathbf{H}\left(\mathbf{D}\right)\boldsymbol{\theta}\right)^{H}\left(\mathbf{x}-\mathbf{H}\left(\mathbf{D}\right)\boldsymbol{\theta}\right)\right]. \quad (6)$$

Because of the condition of white Gaussian noise, ML estimation is equivalent to least square (LS) estimation

$$\left(\hat{\mathbf{D}},\hat{\boldsymbol{\theta}}\right)=\arg\min_{\mathbf{D},\boldsymbol{\theta}}\left[\left(\mathbf{x}-\mathbf{H}\left(\mathbf{D}\right)\boldsymbol{\theta}\right)^{H}\left(\mathbf{x}-\mathbf{H}\left(\mathbf{D}\right)\boldsymbol{\theta}\right)\right]. \quad (7)$$

Note that $\mathbf{x}$ is linearly related to $\boldsymbol{\theta}$ while nonlinearly to $\mathbf{D}$. After some simple derivation, the estimation scheme can be partitioned into two consecutive steps [3, 8–10]. Firstly, estimate $\mathbf{D}$:

$$\hat{\mathbf{D}}=\arg\max_{\mathbf{D}}\left[\mathbf{x}^{H}\left(\mathbf{H}\left(\mathbf{D}\right)\left(\mathbf{H}^{H}\left(\mathbf{D}\right)\mathbf{H}\left(\mathbf{D}\right)\right)^{-1}\mathbf{H}^{H}\left(\mathbf{D}\right)\right)\mathbf{x}\right]. \quad (8)$$

Secondly, calculate the estimate of $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}=\left(\mathbf{H}^{H}\left(\hat{\mathbf{D}}\right)\mathbf{H}\left(\hat{\mathbf{D}}\right)\right)^{-1}\mathbf{H}^{H}\left(\hat{\mathbf{D}}\right)\mathbf{x} \quad (9)$$

and obtain $\hat{A}^{(p)}$ and $\hat{\phi}_0^{(p)}$ just from getting the amplitude and angle of the pth entry of $\hat{\boldsymbol{\theta}}$, respectively.

Since the second step is trivial, we focus on the first one. There are KP parameters to be estimated. For this optimization, an appropriate guess of the initial value is crucial if an iterative algorithm is applied, or the joint search should be conducted on KP dimensions if the grid search method is applied. Motivated by the results of [8] and [9], we develop an importance sampling estimator, which is a global optimizer reducing the computational complexity.

3. IMPORTANCE SAMPLING ESTIMATOR

Define the compressed likelihood function

$$L\left(\mathbf{D}\right)=\exp\left[\rho_{L}\mathbf{x}^{H}\left(\mathbf{H}\left(\mathbf{D}\right)\left(\mathbf{H}^{H}\left(\mathbf{D}\right)\mathbf{H}\left(\mathbf{D}\right)\right)^{-1}\mathbf{H}^{H}\left(\mathbf{D}\right)\right)\mathbf{x}\right] \quad (10)$$

and the corresponding normalized version

$$\bar{L}\left(\mathbf{D}\right)=\frac{L\left(\mathbf{D}\right)}{\int\ldots\int L\left(\mathbf{D}\right)d\mathbf{d}^{(1)}d\mathbf{d}^{(2)}\ldots d\mathbf{d}^{(K)}} \quad (11)$$

which is refer to as a pseudo PDF. Both the compressed likelihood function and the pseudo PDF are positive since $\mathbf{H}^{H}\left(\mathbf{D}\right)\mathbf{H}\left(\mathbf{D}\right)$ satisfies the properties of a positive definite matrix. The maximization of (8) is equivalent to the maximization of the pseudo PDF (11). Thus, based on the theorem [11], we get the estimate

$$\hat{\mathbf{d}}^{(k)} = \int \dots \int \mathbf{d}^{(k)} \bar{L}(\mathbf{D})\, d\mathbf{d}^{(1)} d\mathbf{d}^{(2)} \dots d\mathbf{d}^{(K)}. \tag{12}$$

As reported in [9], the global optimum is attained for $\rho_L \to \infty$ while in practice the value of ρ_L can be chosen to a finite number for a specific problem. Eq. (11) provides a closed-form expression for the estimates; however, this approach also involves multidimensional integration, with the number of dimensions being the same as that of the grid search in ML estimation. In this paper, we use Monte Carlo method to solve the involved integration. Nevertheless, directly generating samples from the pseudo PDF (11) is tricky. Hence, we use importance sampling which generates samples via a simpler PDF. Meanwhile, it alleviates the so-called 'curse of dimensionality'.

Let $\bar{g}(\mathbf{D})$ be another pseudo PDF whose samples are easy to be generated. In importance sampling, the expected value of a function $h(\mathbf{D})$ with respect to $\bar{L}(\mathbf{D})$ could be converted to

$$\begin{aligned} \hat{h}(\mathbf{D}) &= \int \dots \int h(\mathbf{D})\, \bar{L}(\mathbf{D})\, d\mathbf{d}^{(1)} d\mathbf{d}^{(2)} \dots d\mathbf{d}^{(K)} \\ &= \int \dots \int h(\mathbf{D})\, \frac{\bar{L}(\mathbf{D})}{\bar{g}(\mathbf{D})} \bar{g}(\mathbf{D})\, d\mathbf{d}^{(1)} d\mathbf{d}^{(2)} \dots d\mathbf{d}^{(K)}. \end{aligned} \tag{13}$$

Then it is equivalent to the expected value of $h(\mathbf{D}) \left(\bar{L}(\mathbf{D}) / \bar{g}(\mathbf{D})\right)$ with respect to $\bar{g}(\mathbf{D})$, which is also referred to as the normalized importance function. After the samples are drawn, the implementation of (13) using Monte Carlo approximation is

$$\frac{1}{R} \sum_{r=1}^{R} h\left(\mathbf{D}^{(r)}\right) \frac{\bar{L}\left(\mathbf{D}^{(r)}\right)}{\bar{g}\left(\mathbf{D}^{(r)}\right)} \tag{14}$$

where $\mathbf{D}^{(r)}$ is the rth realization of $\mathbf{D}$ generated from the pseudo PDF $\bar{g}(\mathbf{D})$. One issue is the choice of $\bar{g}(\mathbf{D})$. On the one hand, $\bar{g}(\mathbf{D})$ should be simple so that the samples could be easily generated. On the other hand, $\bar{g}(\mathbf{D})$ should be similar to $\bar{L}(\mathbf{D})$ to reduce the variance of the estimates. To satisfy both requirements, considering $\mathbf{H}^H(\mathbf{D})\, \mathbf{H}(\mathbf{D})$ from the pseudo PDF (11) to be an identity matrix multiplied by N, the importance function is chosen as

$$\begin{aligned} g(\mathbf{D}) &= \exp\left[\rho_g \mathbf{x}^H \left(\frac{1}{N} \mathbf{H}(\mathbf{D})\, \mathbf{H}^H(\mathbf{D})\right) \mathbf{x}\right] \\ &= \prod_{p=1}^{P} \exp\left[\rho_g I\left(d^{(p,1)}, d^{(p,2)}, \dots, d^{(p,K)}\right)\right] \end{aligned} \tag{15}$$

where

$$I\left(d^{(p,1)},d^{(p,2)},\ldots,d^{(p,K)}\right)=\frac{1}{N}\left|\sum_{n=0}^{N-1}x(n)\exp\left(-j2\pi\left(\sum_{k=1}^{K}d^{(p,k)}\frac{n^k}{kN^{k-1}}\right)\right)\right|^2 \tag{16}$$

and the normalized version of $g(\mathbf{D})$ is

$$\bar{g}(\mathbf{D})=\frac{g(\mathbf{D})}{\int\ldots\int g(\mathbf{D})\,d\mathbf{d}^{(1)}d\mathbf{d}^{(2)}\ldots d\mathbf{d}^{(K)}}=\prod_{p=1}^{P}\bar{g}\left(d^{(p,1)},d^{(p,2)},\ldots,d^{(p,K)}\right) \tag{17}$$

where $\bar{g}(d^{(p,1)},d^{(p,2)},\ldots,d^{(p,K)})$ is the normalized version of $\exp[\rho_g I(d^{(p,1)}, d^{(p,2)},\ldots,d^{(p,K)})]$.

Thus, in a sense, the importance function alleviates the effect of cross-term, which is crucial in the analysis of multicomponents. In theory, a moderate value ρ_g is advisable. On the one hand, one component may overwhelm others if ρ_g is too large. On the other hand, the influence of noise may be magnified to the level of the signals if ρ_g is too small. However, as reported in [9], ρ_g is not a very sensitive parameter.

4. STEPS OF ALGORITHM

Based on derivation in Section 3, the steps to estimate the PCs are as follows:

Step 1. Choose a large number M and let $\delta = 1/M$ which determines the resolution. Uniformly divide the range $(0,1)$ into M grids, and calculate the importance function from the data at M^K discrete points in a K dimension space. Hence, we obtain each value of the importance function $\bar{g}(d_{(m_1)},d_{(m_2)},\ldots,d_{(m_K)})$ corresponding to the joint coordinate $(d_{(m_1)},d_{(m_2)},\ldots,d_{(m_K)})$ where $m_k = 1,2,\ldots,M$ and $d_{(m_k)} = (m_k-1)\delta$ for $k=1,2,\ldots,K$.

Step 2. Calculate the marginal PDF:

$$\bar{g}\left(d_{(m_1)}\right)=\sum_{m_2=1}^{M}\sum_{m_3=1}^{M}\cdots\sum_{m_K=1}^{M}\bar{g}\left(d_{(m_1)},d_{(m_2)},\ldots,d_{(m_K)}\right)\delta^{K-1} \tag{18}$$

and obtain the cumulative distribution function:

$$G\left(d_{(m_1)}\right)=\sum_{m=1}^{m_1}\bar{g}\left(d_{(m)}\right)\delta \tag{19}$$

Step 3. Calculate the conditional PDF:

$$\bar{g}\left(d_{(m_2)}/d_{(m_1)}\right) = \frac{\bar{g}\left(d_{(m_1)}, d_{(m_2)}\right)}{\bar{g}\left(d_{(m_1)}\right)} \tag{20}$$

where the numerator is calculated by

$$\bar{g}\left(d_{(m_1)}, d_{(m_2)}\right) = \sum_{m_3=1}^{M}\sum_{m_4=1}^{M}\cdots\sum_{m_K=1}^{M}\bar{g}\left(d_{(m_1)}, d_{(m_2)}, \ldots, d_{(m_K)}\right)\delta^{K-2} \tag{21}$$

and obtain the conditional cumulative distribution function

$$G\left(d_{(m_2)}/d_{(m_1)}\right) = \sum_{m=1}^{m_2}\bar{g}\left(d_{(m)}/d_{(m_1)}\right)\delta. \tag{22}$$

Similarly, calculate the conditional PDF

$$\bar{g}\left(d_{(m_3)}/d_{(m_1)}, d_{(m_2)}\right) = \frac{\bar{g}\left(d_{(m_1)}, d_{(m_2)}, d_{(m_3)}\right)}{\bar{g}\left(d_{(m_1)}, d_{(m_2)}\right)} \tag{23}$$

where the numerator is calculated by

$$\bar{g}\left(d_{(m_1)}, d_{(m_2)}, d_{(m_3)}\right) = \sum_{m_4=1}^{M}\sum_{m_5=1}^{M}\cdots\sum_{m_K=1}^{M}\bar{g}\left(d_{(m_1)}, d_{(m_2)}, \ldots, d_{(m_K)}\right)\delta^{K-3} \tag{24}$$

and obtain the conditional cumulative distribution function

$$G\left(d_{(m_3)}/d_{(m_1)}, d_{(m_2)}\right) = \sum_{m=1}^{m_3}\bar{g}\left(d_{(m)}/d_{(m_1)}, d_{(m_2)}\right)\delta. \tag{25}$$

Then, repeat similar calculations until $\bar{g}(d_{(m_K)}/d_{(m_1)}, d_{(m_2)}, \ldots, d_{(m_{K-1})})$ and $G(d_{(m_K)}/d_{(m_1)}, d_{(m_2)}, \ldots, d_{(m_{K-1})})$ are obtained.

Step 4. For an index $r = 1, 2, \ldots, R$, generate random numbers $u^{(1,1)} \sim U[0,1], u^{(1,2)} \sim U[0,1], \ldots, u^{(1,K)} \sim U[0,1]$ for the first component, where $U[0,1]$ denotes the uniform distribution ranging from 0 to 1, and obtain

$$\begin{aligned}
d^{(1,1)(r)} &= \arg\min_{d_{(m_1)}}\left|u^{(1,1)} - G\left(d_{(m_1)}\right)\right| \\
d^{(1,2)(r)} &= \arg\min_{d_{(m_2)}}\left|u^{(1,2)} - G\left(d_{(m_2)}/d^{(1,1)(r)}\right)\right| \\
&\cdots \\
d^{(1,K)(r)} &= \arg\min_{d_{(m_K)}}\left|u^{(1,K)} - G\left(d_{(m_K)}/d^{(1,1)(r)}, d^{(1,2)(r)}, \ldots, d^{(1,K-1)(r)}\right)\right|
\end{aligned} \tag{26}$$

one by one. Similarly, get $d^{(p,1)(r)}, d^{(p,2)(r)}, \ldots, d^{(p,K)(r)}$, $p=2,3,\ldots,P$ for the pth component. It should be noted that at least one parameter in the different parameter pairs $\{d^{(p,1)}, d^{(p,2)}, \ldots, d^{(p,K)}\}$ is different from the others. Subsequently, reorder components by arranging the first order PCs $d^{(1,1)(r)}, d^{(2,1)(r)}, \ldots, d^{(P,1)(r)}$ in ascending order. Consequently, we get $\mathbf{d}^{(k)(r)} = [d^{(1,k)(r)}, d^{(2,k)(r)}, \ldots, d^{(P,k)(r)}]^T$ for $k = 1, 2, \ldots, K$.

Step 5. After R realizations are completed, calculate the circular mean which replaces the linear mean to reduce the computations. The estimate of the PC using the circular mean is given by

$$\hat{\mathbf{d}}^{(k)} = \frac{1}{2\pi}\angle\frac{1}{R}\sum_{r=1}^{R}\frac{L\left(\mathbf{D}^{(r)}\right)}{g\left(\mathbf{D}^{(r)}\right)}\exp\left(j2\pi\mathbf{d}^{(k)(r)}\right) \tag{27}$$

where the mark $\angle$ denotes getting the angle of a complex number, and $L(\mathbf{D}^{(r)})/g(\mathbf{D}^{(r)})$ is the importance weight. Note that there is no normalized operation for both $L(\mathbf{D}^{(r)})$ and $g(\mathbf{D}^{(r)})$, since either normalized number is a constant for any r and can be ignored in the computation of the angle.

Briefly, the computational complexity of this algorithm consists of two parts. The first one is from the calculation of $\bar{g}\left(d_{(m_1)}, d_{(m_2)}, \ldots, d_{(m_K)}\right)$ as well as other conditional PDFs in a K dimension space, which is common for all components, while the second one is from the generation of samples, which is linearly related to the number of components. As a comparison, if the grid search is directly applied in ML estimation, the computational complexity is exponentially related to KP.

5. CRLBS OF MULTICOMPONENT PPSS

In the case that the amplitudes and initial phases are unknown, References [12, 13] give the CRLBs of multicomponent signals. Using our notations, the $KP \times KP$ Fisher information matrix is

$$\mathbf{F} = \frac{2}{\sigma^2}\mathrm{Re}\left[\left(\mathbf{U}^H\left(\mathbf{I}_{N\times N} - \mathbf{H}\left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\right)\mathbf{U}\right)\odot\left(\mathbf{Q}^T\otimes\mathbf{1}_{K\times K}\right)\right] \tag{28}$$

where the notations are as follows:

- $\odot$ Schur-Hadamard product
- $\otimes$ Kronecker product
- $\mathbf{I}_{N\times N}$ $N \times N$ identity matrix
- $\mathbf{1}_{K\times K}$ $K \times K$ matrix of ones
- $\mathbf{U} = \left[\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(P)}\right]$ $N \times KP$ matrix

$$\mathbf{U}^{(p)} = \left[\frac{\partial \mathbf{h}^{(p)}}{\partial d^{(p,1)}}, \frac{\partial \mathbf{h}^{(p)}}{\partial d^{(p,2)}}, \ldots, \frac{\partial \mathbf{h}^{(p)}}{\partial d^{(p,K)}}\right] \quad N \times K \text{ matrix}$$
$$\mathbf{Q} = \boldsymbol{\theta}\boldsymbol{\theta}^H \quad P \times P \text{ matrix}$$

Besides, see Section 2 for the definitions of $\mathbf{H}$, $\mathbf{h}$, and $\boldsymbol{\theta}$. For clarity, $\mathbf{H}(\mathbf{D})$ and $\mathbf{h}(\mathbf{D})$ in Section 2 are rewritten as $\mathbf{H}$ and $\mathbf{h}$, respectively.

By using the Fisher information matrix, the CRLB is defined as

$$\text{CRLB}_i = \left[\mathbf{F}^{-1}\right]_{ii}$$

where $[\cdot]_{ii}$ denotes getting the ith diagonal entry. The order of the corresponding PCs is $\{d^{(1,1)}, d^{(1,2)}, \ldots, d^{(1,K)}, \ldots, d^{(P,1)}, d^{(P,2)}, \ldots, d^{(P,K)}\}$. Specifically, [3] proves that in the case of only two PPSs, if the amplitudes of the two components are the same, the CRLBs of the corresponding PCs are the same, i.e.,

$$\text{CRLB}\left(d^{(1,k)}\right) = \text{CRLB}\left(d^{(2,k)}\right).$$

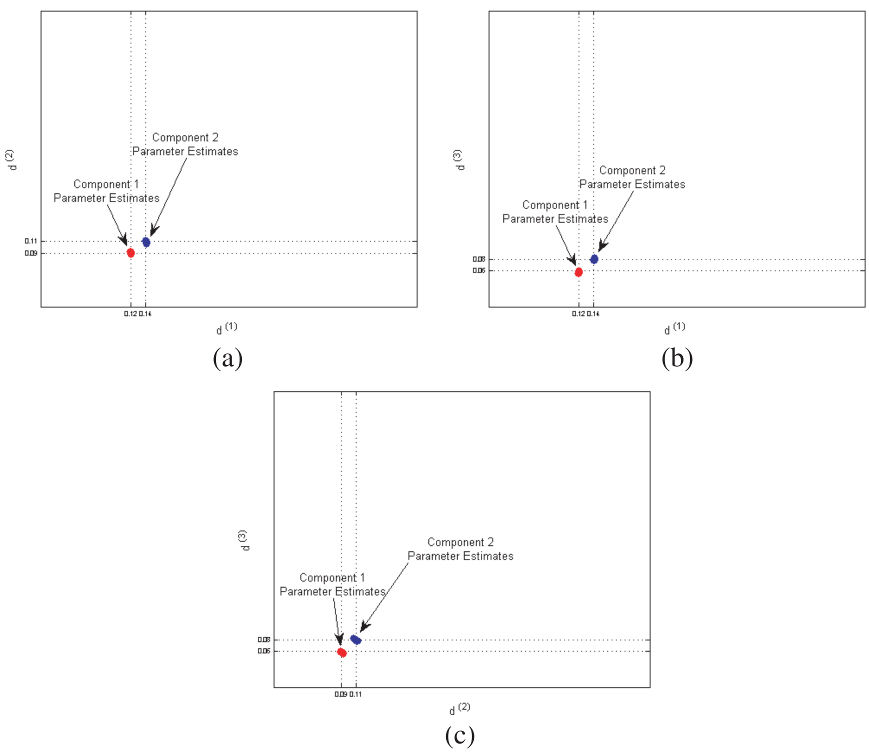
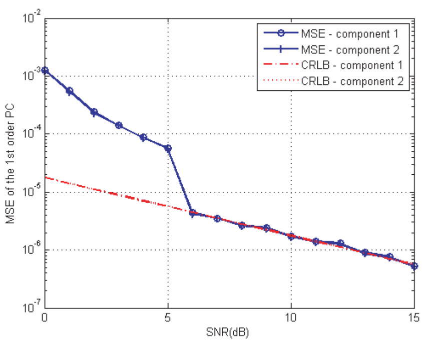


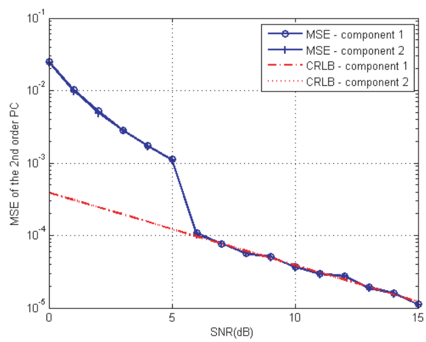
Figure 1. (a) Scatter diagram of estimates over the plane of $(d^{(1)}, d^{(2)})$ plane; (b) Scatter diagram of estimates over the plane of $(d^{(1)}, d^{(3)})$; (c) Scatter diagram of estimates over the plane of $(d^{(2)}, d^{(3)})$.

6. EXPERIMENTS AND RESULTS

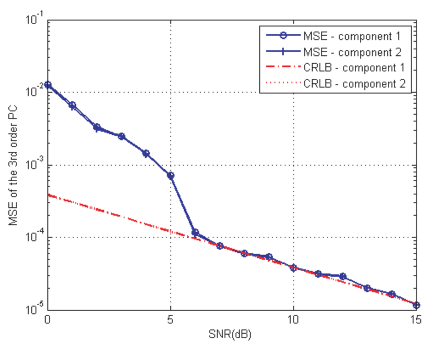
To illustrate the algorithm's performance, two closely spaced quadratic frequency modulated signals are used as components in our simulation. Similar to [8, 9], in the following experiments, components are equi-amplitude with $A^{(1)} = A^{(2)} = 1$. The initial phases are $\phi_0^{(1)} = 0$ and $\phi_0^{(2)} = \pi/4$, respectively, and the PC sets are $\left\{d^{(1,1)}, d^{(1,2)}, d^{(1,3)}\right\} = \{0.12, 0.09, 0.06\}$ and $\left\{d^{(2,1)}, d^{(2,2)}, d^{(2,3)}\right\} = \{0.14, 0.11, 0.08\}$, respectively. The sample length is 100. In experiments, the parameters δ and R are set to be 0.001 and 5000, respectively. Besides, as mentioned above, the choice of ρ_L and ρ_g are problem specific. In our simulation, they are set to be 2 and 0.2, respectively. Using the first component, the signal-to-noise ratio (SNR) is defined as $10 \lg \left(A^{(1)2}/\sigma^2\right)$.



(a)



(b)



(c)

Figure 2. (a) MSE and CRLB of the 1st order PC versus SNR; (b) MSE and CRLB of the 2nd order PC versus SNR; (c) MSE and CRLB of the 3rd order PC versus SNR.

Experiment 1 simulates the estimation of PCs where the SNR is chosen as 10 dB. In Figs. 1(a), (b) and (c), we scatter the projection of estimates over the plane of $\left(d^{(1)}, d^{(2)}\right)$, the one of $\left(d^{(1)}, d^{(3)}\right)$, and the one of $\left(d^{(2)}, d^{(3)}\right)$, respectively. The estimates are shown by circles. 100 trials for each component are plotted. It can be shown that the estimates are close to the true parameters in the case of closely spaced signals.

To further evaluate the performance of this algorithm, Experiment 2 gives the Monte Carlo simulation results which are shown using mean square error (MSE) from the average of 500 Monte Carlo simulations. As a comparison, the CRLBs of the PCs are used. Figs. 2(a), (b) and (c) show the estimation performance of the three coefficients, respectively. Note that the lines of the CRLBs may be different from those of other literatures since we adopt the normalized scale of PCs, which is shown in Section 2. It can be observed that the CRLB is attained when the SNR reaches 6 dB.

7. CONCLUSION

In this paper, we developed a ML parameter estimation algorithm of multicomponent PPSs using importance sampling. The ML estimation scheme was partitioned to two consecutive steps to separately estimate the nonlinear parameter set including the PCs and the linear one including the amplitudes and the initial phases. Then aiming at the former estimation, we utilized importance sampling to alleviate the effect of the so-called cross-term between components. In a ML estimation sense, it is a non-iterative algorithm avoiding the choice of the initial guess which is somewhat tricky. However, it should be noted that with the increase of the number of polynomial orders, further efficient algorithm will need to be investigated.

REFERENCES

1. Abatzoglou, T. J., "Fast maximum likelihood joint estimation of frequency and frequency rate," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 22, No. 6, 708–715, 1986.
2. Angeby, J., "Estimating signal parameters using the nonlinear instantaneous least squares approach," *IEEE Transactions on Signal Processing,* Vol. 48, No. 10, 2721–2732, 2000.
3. Pham, D. S. and A. M. Zoubir, "Analysis of multicomponent polynomial phase signals," *IEEE Transactions on Signal Processing,* Vol. 55, No. 1, 56–65, 2007.

4. Peleg, S. and B. Porat, "Estimation and classification of polynomial-phase signals," *IEEE Transactions on Information Theory,* Vol. 37, No. 2, 422–430, 1991.
5. Peleg, S. and B. Friedlander, "The discrete polynomial-phase transform," *IEEE Transactions on Signal Processing,* Vol. 43, No. 8, 1901–1914, 1995.
6. Porat, B. and B. Friedlander, "Asymptotic statistical analysis of the high-order ambiguity function for parameter estimation of polynomial-phase signals," *IEEE Transactions on Information Theory,* Vol. 42, No. 3, 995–1001, 1996.
7. Peleg, S. and B. Friedlander, "Multicomponent signal analysis using the polynomial-phase transform," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 32, No. 1, 378–387, 1996.
8. Kay, S. and S. Saha, "Mean likelihood frequency estimation," *IEEE Transactions on Signal Processing,* Vol. 48, No. 7, 1937–1946, 2000.
9. Saha, S. and S. M. Kay, "Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling," *IEEE Transactions on Signal Processing,* Vol. 50, No. 2, 224–230, 2002.
10. Kay, S. M., *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory,* 1993.
11. Pincus, M., "A closed form solution for certain programming problems," *Oper. Res.*, 690–694, 1962.
12. Stoica, P. and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing,* Vol. 37, No. 5, 720–741, 1989.
13. Yau, S. F. and Y. Bresler, "A compact Cramer-Rao bound expression for parametric estimation of superimposed signals," *IEEE Transactions on Signal Processing,* Vol. 40, No. 5, 1226–1230, 1992.