Extended Kalman Filter for Parameter Estimation of Wideband Polynomial Phase Signals in Sensors Arrays

A. OULDALI, S. SADOUDI, Z. MESSAOUDI
Laboratory of Communication Systems
Military Polytechnic School
B.P 17, Bordj El Bahri, Algiers
ALGERIA

Abstract: - In this paper, we consider the problem of estimating wideband polynomial phase signals in sensor arrays. The estimation is based on the introduction of an exact but nonlinear state space modelization of the signal, which compels us to use the extended Kalman filter (EKF). Furthermore, a solution to the problem of initialization of the filter is also proposed. Under this solution, the numerical simulations show that the use of the EKF improves existing methods in terms of statistical performances since the EKF-based estimators exhibit high performances.

Key-Words: - Cramer-Rao bound, Extended Kalman filter, Wideband polynomial phase signals.

1 Introduction
Parameter estimation of polynomial phase signal (PPS) is an important problem in many engineering applications such as radar and communications. Recently, there has been a growing interest in estimating wideband PPS impinging on a sensor arrays [1-6]. In [3], a new form of the maximum likelihood estimator of signal parameters is presented. However, since the proposed estimator is computationally intensive an approximate technique (the chirp beamformer) is proposed. In [4], the application of the high-order instantaneous moment transforms the PPS array signal into stationary joint angle-frequency estimation (JAFE) problem which is based on ESPRIT algorithm [4] [7]. In fact, it has been shown in [4] that it is possible to jointly estimate the highest order phase coefficient (HOC) and the direction of arrival (DOA) of the signal. In the following, we call this approach “joint angle highest order coefficient estimation” (JAHOCE). In [1-2] [5-6], the approaches suffer from problems. In fact, in [1-2] the method is restricted by short sliding data window lengths. In [5], the approach assumes linear FM signals with known central frequency. In [6], the iterative approach may lead to strongly biased DOA estimates [1] and its convergence is not guaranteed [3].

In this paper, we address the problem of estimating wideband PPS affected by additive noise. The principle of estimation is based on the introduction of an exact but nonlinear state model, of the PPS, which compels us to use the extended Kalman filter (EKF) which has already been used for estimating narrowband PPS [8] but which has never been used in the case of wideband PPS to the best of our knowledge. Moreover, we propose a solution to the problem of initialization of the filter since the initial conditions of the state model are assumed to be unavailable. Under this solution, the numerical simulations show that the EKF improves JAHOCE in terms of statistical performances. In fact the EKF-based estimators exhibit high performances since our proposed method exploits implicitly the double of the initial number of snapshots.

The present paper is organized as follows. The array signal model used in this paper is presented in Section 2. In section 3, we introduce the state space modelization of the PPS. Then, section 4 presents the EKF-based estimators. We present in section 5 the solution to the problem of initialization of the filter and the numerical results are given in section 6. Finally, section 7 concludes this paper.

2 Array signal model
Let a wideband PPS \( s(n) \) impinges, from an unknown DOA \( \theta \), on a uniform linear array (ULA) antenna of \( L \) sensors. Then, the \( L \times 1 \) vector array outputs is given by [3-4]

\[
y(n) = a(\theta, n) s(n) + w(n), \quad n = 0, \ldots, Ne - 1
\]  

where \( a(\theta, n) \) is the \( L \times 1 \) time-varying steering vector, \( s(n) \) is the PPS waveform, \( w(n) \) is the \( L \times 1 \) vector of complex circularly Gaussian zero-mean temporally and spatially white noise with known variance \( \sigma^2 \) and \( Ne \) is the number of snapshots.

The PPS \( s(n) \) is given by

\[
s(n) = A \exp \left\{ \sum_{i=0}^{N} a_i (n\Delta)^i \right\}
\]  

(2)
where $A > 0$, $\{a_i\}_{i=0,...,N}$ are the phase coefficients, $N$ is the degree of the polynomial phase and $\Delta$ is the sampling period. In the following, $N$ is assumed to be known.

The time-varying steering vector $\mathbf{a}(\theta, n)$ is given by [3-4] 
$$
\mathbf{a}(\theta, n) = \left[1, e^{j\theta}, \ldots, e^{jL\theta}\right]^T
$$
(3)
where 
$$
f(n) = \sum_{i=0}^{N-1} (i+1) a_i(n\Delta)^i
$$
(4)
is the instantaneous frequency, of the signal, assumed to be constant during the time necessary for a wave to travel across the array aperture and 
$$
\psi = \frac{d}{c}\sin(\theta)
$$
(5)
where $d$ is the spacing between two adjacent sensors and $c$ is the propagation speed in the medium.

3 State space modelization

Let $x(n)$ be the vector of the unknown parameters to be estimate 
$$
x(n) = [A, \theta, a_0, \ldots, a_N]^T
$$
(6)
Thanks to this vector, we can represent the signal (1) by the following exact but nonlinear state model characterized, fortunately, by an evolution matrix equals to identity 
$$
x(n+1) = x(n)
$$
$$
y_{1,1}(n) = g_{1,1}(x(n)) + \text{Re}\{w_1(n)\}
$$
$$
y_{1,2}(n) = g_{1,2}(x(n)) + \text{Im}\{w_1(n)\}
$$
$$
\vdots
$$
y_{L,1}(n) = g_{L,1}(x(n)) + \text{Re}\{w_L(n)\}
$$
y_{L,2}(n) = g_{L,2}(x(n)) + \text{Im}\{w_L(n)\}
(7)
where $y_{1,1}(n) = \text{Re}\{y(n)\}$, $y_{1,2}(n) = \text{Im}\{y(n)\}$ (with $y(n)$ and $w(n)$) are the $l$th row of $y(n)$ and $w(n)$ and 
$$
g_{1,1}(x(n)) = x_1(n)\cos[\Phi_1(x(n))]
$$
(8)
g_{1,2}(x(n)) = x_1(n)\sin[\Phi_1(x(n))]
(9)
$$
\Phi_1(x(n)) = x_{N+1}(n) + \sum_{i=0}^{N-1} x_{i+1}(n) + (l-1)
\]$$
x((i+1)x_{i+1}(n)\left[\frac{d}{c}\sin(x_2(n))\right](n\Delta)^i)
$$
(10)

4 Extended Kalman filter based estimators

In order to obtain the unknown parameters $\{A, \theta, a_0, \ldots, a_N\}$, we should estimate the state vector $x(n)$ thanks to the EKF since the proposed state model is nonlinear. The following algorithm summarizes the EKF and proposes the EKF-based estimators, of the unknown parameters, given $N_e$ noisy observations $y(n)$.

**Algorithm 1**

1) Initial Conditions ($n = 0$) 
$$
\begin{align*}
\hat{x}_p(0) &= E[x(0)] \\
\hat{P}_p(0) &= E\left[(x(0) - \hat{x}_p(0)) (x(0) - \hat{x}_p(0))^T\right]
\end{align*}
$$
(11)
(12)
2) Update equations 
$$
\begin{align*}
K(n) &= \hat{P}_p(n)J^T(\hat{z}_p(n))J(\hat{z}_p(n))\hat{P}_p(n) + \frac{\sigma^2}{2}I \\
\hat{y}_1(n) &= y_{1,1}(n) - g_{1,1}(\hat{x}_p(n)) \\
\vdots \\
\hat{y}_L(n) &= y_{L,1}(n) - g_{L,1}(\hat{x}_p(n)) \\
\hat{y}(n) &= [\hat{y}_1(n), \hat{y}_2(n), \ldots, \hat{y}_L(n)]^T \\
\hat{z}(n) &= \hat{x}_p(n) + K(n) \hat{y}(n) \\
P(n) &= \hat{P}_p(n) - K(n)J(\hat{z}_p(n))\hat{P}_p(n)
\end{align*}
$$
(13)
(14)
(15)
(16)
(17)
3) Prediction equations 
$$
\begin{align*}
\hat{x}_p(n+1) &= \hat{x}(n) \\
\hat{P}_p(n+1) &= P(n)
\end{align*}
$$
(18)
(19)
4) $n = n + 1$. If $n \leq N_e - 1$ go to step 2)

5) EKF-based estimators 
$$
\begin{align*}
\hat{A}, \hat{\theta}, \hat{a}_0, \ldots, \hat{a}_N
\end{align*}
$$
(20)
where in particular $I$ is the $(2L \times 2L)$ identity matrix, $x(\bullet)$ is the estimated state vector, $K(n)$ is the Kalman gain and $J(\bullet)$ is the Jacobean of $\{g_{1,1}(\bullet), g_{1,2}(\bullet), \ldots, g_{L,1}(\bullet), g_{L,2}(\bullet)\}$.

5 Solution to the problem of initialization

Before to solve the problem of initialization of the EKF, we should

- See the effect of demodulation of the output of each sensor after obtaining the DOA $\theta$ and the HOC $a_N$ thanks to JAHOCE proposed in [4].
- Recall the high-order ambiguity function (HAF) generally used to estimate PPS.
- Present the Fisher information matrix (FIM) which is necessary to obtain the CRB of each parameter.

5.1 Effect of demodulation

Let $z(n) = y(n) - w(n)$. The $l$th row of $z(n)$ is given by 
$$
z_{l}(n) = A \exp\left(\int_{-N_e}(n\Delta)^l\right)
$$
(10)
\[
\sum_{i=0}^{N-1} \left( a_i + (i+1)a_i(l-1)\psi(n\Delta)^i \right) \right]\]  
\[
\exp \left\{ -j [n\Delta + N(l-1)\psi] a_N (n\Delta)^{N-1} \right\}
\]
leads to the following PPS
\[
x(n) = A \exp \left\{ j \left( \sum_{i=0}^{N-1} a_i (n\Delta)^i + \sum_{i=0}^{N-2} (i+1)a_i(l-1)\psi(n\Delta)^i \right) \right\}
\]
which is characterized by a polynomial phase of degree \(N-1\) and with a HOC \(a_{N-1}\).

### 5.2 High-order ambiguity function

The HAF, of order \(Q\), of a signal \(v(n)\) is defined as follows [9]
\[
P_Q(v, \theta, \tau) = \sum_{n=0}^{N-1} \sum_{\tau=0}^{Q-1} I_{k} \{ v(n-k\tau) \} \exp\left\{ j n\Delta (n\Delta)^Q \right\}
\]
where \(I_k \{ v(n) \} = v(n), I_{k+1} \{ v(n) \} = \psi(v(n))\) and \(h_{Q,k} = \binom{Q}{k}\).

#### Main property:
For \(x(n)\) given by (23), we have [9]
\[
a_{N-1} = \frac{1}{(N-1)!(\Delta)^{N-2}} \arg \max_\theta \{ P_{N-1}(x, \theta, \tau) \}
\]
Thus the HOC \(a_{N-1}\) of the PPS (23), can be obtained from the peak of \(P_{N-1}(x, \theta, \tau)\).

### 5.3 Cramer-Rao Bounds

Using the properties of the additive noise \(w(n)\), it is possible to show that the FIM, noted \(\mathcal{F}\), of \(\mathbf{P} = [A, \theta, a_1, \ldots, a_N]^T\), is given by
\[
\mathcal{F} = \begin{bmatrix}
\frac{2NeL}{\sigma^2} & 0 \\
0 & \frac{2A^2}{\sigma^2}
\end{bmatrix}
\]
where \(R = \sum_{n=0}^{N-1} G_2 \{ f(n) \} \frac{d\psi}{d\theta} \) \(S_j = \sum_{n=0}^{N-1} (n\Delta)^j \{ (n\Delta)G_1 + jG_2 \psi \} f(n) \frac{d\psi}{d\theta} \)
\(T_{k,j} = \sum_{n=0}^{N-1} \{ i+j \} (n\Delta)^j \{ (n\Delta)G_2 \psi + L(n\Delta)^2 \} (n\Delta)^j \)
where \(f(n)\) is given by (4), \(\psi\) is given by (5) and
\[
G_k = \sum_{m=0}^{L-1} m^k
\]
The CRB's of \(P\) are given by
\[
CRB(P_j) = \left( \mathcal{F}^{-1} \right)_{j,j} \]

### 5.4 Solution to the problem of initialization

#### Algorithm 2

1) Estimation step
1.a) Estimate \(a_N\) and \(\theta\) thanks to JAHCOE of [4]
1.b) Let \(l = 1\) and \(\hat{\psi} = d \sin(\hat{\theta})/c\)
1.c) Let \(\hat{a}_{l,N} = \hat{a}_{N} \quad m = N, \quad y_{l}^{(m)}(n) = y_{l}(n)\) and do
\[
y_{l}^{(m-1)}(n) = y_{l}^{(m)}(n) \exp\left\{ -j [n\Delta + m(l-1)\psi] \hat{a}_{m}(n\Delta)^{m-1} \right\}
\]
1.d) Choose \(\tau = Ne/(m-1)\) [9] and do
\[
\hat{a}_{l,m-1} = \frac{1}{(m-1)!(\Delta)^{m-2}} \arg \max_\theta \{ P_{m-1}(y_{l}^{(m-1)}(n), \theta, \tau) \}
\]
1.e) \(m = m - 1\). If \(m > 1\) go to step 1.d) else do
\[
\hat{a}_{l,0} = \arg \frac{1}{Ne} \sum_{n=0}^{N-1} y_{l}^{(0)}(n)
\]
1.f) \(l = l + 1\). If \(l \leq L\) go to step 1.c) else do
\[
\hat{A} = \frac{1}{L} \sum_{i=1}^{L} \hat{A}_i
\]
2) Initialization step
\[
\hat{A}_0(0) = \left[ \hat{A}, \hat{\theta}, \hat{a}_0, \ldots, \hat{a}_N \right]^T
\]
\[
P_0(0) = \text{diag} \begin{bmatrix} \lambda_1 \var{\hat{\lambda}_1}, \lambda_2 \var{\hat{\lambda}_2}, \ldots, \lambda_N \var{\hat{\lambda}_N} \end{bmatrix}
\]
where \(\text{diag} \{\bullet\}\) denotes the diagonal matrix, \(\var{\bullet}\) is the theoretical variance (TV) of the corresponding estimator and \(\{\lambda_1, \ldots, \lambda_N\} > 0\).
From (40), we see that the initial matrix \(P_0(0)\) needs the TV of each estimator. However, these TV are too much difficult to derive and this point is beyond the scope of this paper. To overcome this problem, we can choose \(P_0(0)\) according to the following equation
\[
P_0(0) = \text{diag} \begin{bmatrix} \mu_{rcrb} \var{\hat{\mu}_{rcrb}}, \mu_{crb} \var{\hat{\mu}_{crb}}, \var{\hat{\mu}_{crb}} \end{bmatrix}
\]
\[ \mu \text{crb}\{a_1\}, \ldots, \mu_{N+i}\text{crb}\{a_N\} \]  

(41)

where \( \text{crb}\{\bullet\} \) denotes the CRB of the corresponding parameter and \( \{\kappa_1, \ldots, \kappa_{i+3}\} \geq 1 \). \( \mu_i = 1 \) (respectively \( \mu_i >> 1 \)) means that the \( i \) th component of \( x_0(0) \) is efficient (respectively is far from the true value to be estimated). In order to get an idea about the coefficient \( \mu_i \), we can evaluate the ratio \( \kappa_i \) of the empirical variance (EV), of the \( i \) th estimator, to its corresponding CRB then we can choose \( \mu_i = \alpha_i \kappa_i \) with \( \alpha_i \geq 1 \) since the EV is not necessary equal to the TV and also with \( \alpha_i \leq B \) (B is a small positive number) to guarantee that

\[ \left[ x(0) - \hat{x}_p(0) \right]^T \left[ x(0) - \hat{x}_p(0) \right] \]  

is of the same order than

\[ E \left[ \left[ x(0) - \hat{x}_p(0) \right]^T \left[ x(0) - \hat{x}_p(0) \right] \right] \]  

(42)

Remark 1: In algorithm 2, the total number of available temporal snapshots, has essentially doubled from \( Ne \) to \( 2Ne \). In fact, the estimation step uses \( Ne \) snapshots, to obtain the initial state vector \( x_j(0) \), and the EKF operates on the same \( Ne \) snapshots. We should emphasize that this data extension (off-line method) will give a significant improvement of the estimation of the unknown parameters \( \{A, \theta, \alpha_0, \ldots, \alpha_N\} \) as it will be seen in the next section.

Remark 2: From (26), we see that \( \text{crb}\{A\} \) is independent from the unknown parameters (the variance \( s^2 \), of the additive noise, is assumed to be known), whereas the others \( \text{crb}\{\bullet\} \) can depend on the unknown values of the parameters. To overcome this problem we can replace these unknown values by their estimate in the CRB expressions.

Remark 3: For on-line estimation, we apply the estimation step of algorithm 2 on the first \( Nr \) snapshots, to obtain the initial state vector \( x_j(Nr) \), then we start the EKF from \( n = Nr \) with \( P_p(Nr) = \text{diag}\{\mu_1\text{crb}\{A\}, \mu_2\text{crb}\{\theta\}, \mu_3\text{crb}\{a_0\}, \ldots, \mu_{N+3}\text{crb}\{a_N\}\} \) where \( \text{crb}\{\bullet\} \) is the CRB evaluated on the first \( Nr \) snapshots.

6 Numerical example

In this section, we begin first by the evaluation of the statistical performances of the estimation step (algorithm 2) in order to obtain the ratios \( \{\kappa_1, \ldots, \kappa_{i+3}\} \). Then, we present the statistical performances of the EKF-based estimators. The SNR is defined as \( \text{SNR} = A^2 / \sigma^2 \).

For the numerical simulations, we consider a quadratic phase signal (QPS) with the following parameters: \( A = 1, N = 2, \Delta = 0.004, a_0 = 0.2\pi, a_1 = 400\pi, a_3 = 200\pi \). The number of snapshots is \( Ne = 256 \), the number of sensors is \( L = 10 \), the DOA of the signal is \( \theta = 40^\circ \), the inter-sensor spacing is \( d = 1.5 \) and \( c = 1500 \).

Remark 4: In the estimation step (algorithm 2), we incorporate in JAHOCE of [4], only temporal smoothing with the \( m \)-factor temporal smoothing \( m = 32 \).

6.1 Statistical performances of estimation step(algorithm2)

In order to evaluate the ratios \( \{\kappa_1, \ldots, \kappa_5\} \) we evaluate the EV of the estimators \( \theta \) and \( a_2 \) (obtained by JAHOCE of [4]) and the estimators (37) and (38) by carrying out 1000 independent realizations of Monte-Carlo.

The following table summarizes some values of \( \{\kappa_1, \ldots, \kappa_5\} \). It shows that the estimator (38), of the amplitude \( A \), is efficient from \( \text{SNR} \geq 0 \) whereas the other estimators, in particular \( \theta \), are not efficient even for high SNR.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>( \kappa_1 )</th>
<th>( \kappa_2 )</th>
<th>( \kappa_3 )</th>
<th>( \kappa_4 )</th>
<th>( \kappa_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>156.4</td>
<td>79.4</td>
<td>2.9</td>
<td>2.23</td>
</tr>
<tr>
<td>2</td>
<td>1.03</td>
<td>121.6</td>
<td>18.1</td>
<td>2.47</td>
<td>1.77</td>
</tr>
<tr>
<td>4</td>
<td>1.05</td>
<td>116.4</td>
<td>14.6</td>
<td>2.19</td>
<td>1.61</td>
</tr>
<tr>
<td>6</td>
<td>1.01</td>
<td>93.2</td>
<td>10.8</td>
<td>1.94</td>
<td>1.55</td>
</tr>
<tr>
<td>8</td>
<td>1.01</td>
<td>103.9</td>
<td>11.9</td>
<td>2.08</td>
<td>1.54</td>
</tr>
<tr>
<td>10</td>
<td>1.03</td>
<td>98.3</td>
<td>11.7</td>
<td>1.91</td>
<td>1.36</td>
</tr>
<tr>
<td>12</td>
<td>1.07</td>
<td>94.6</td>
<td>10.6</td>
<td>2.1</td>
<td>1.45</td>
</tr>
<tr>
<td>14</td>
<td>1.03</td>
<td>88.6</td>
<td>10.4</td>
<td>1.78</td>
<td>1.23</td>
</tr>
<tr>
<td>16</td>
<td>1.03</td>
<td>91.5</td>
<td>10.8</td>
<td>1.73</td>
<td>1.19</td>
</tr>
<tr>
<td>20</td>
<td>1.04</td>
<td>85.5</td>
<td>12.3</td>
<td>1.94</td>
<td>1.3</td>
</tr>
<tr>
<td>25</td>
<td>1.01</td>
<td>93.2</td>
<td>19.3</td>
<td>1.77</td>
<td>1.29</td>
</tr>
</tbody>
</table>

Table 1. Ratio of the empirical variance to CRB

6.2 Statistical performances of the EKF-based estimators

In order to evaluate the statistical performances of the proposed EKF-based estimators, we vary the SNR from 0 dB to 25 dB by step of 2.5 dB and carry out 1000 independent realizations of Monte-Carlo. For each SNR we choose, in equation (41), \( \{\kappa_1 = 1.5\kappa_2\}_1^{-i+3} \).

In all the figures, the CRB’s appear in solid lines, whereas the EV’s, of the EKF-based estimators, are represented by (- -) and the EKVs, of the estimators obtained in the estimation step (algorithm 2), are represented by (- *).

From all the figures, the obtained results show that the EKF-based estimators exhibit performances which reach the CRB for low enough SNR. Thus, our method outperforms, in terms of statistical performances, JAHOCE proposed in [4] for only the estimation of the DOA \( \theta \) and the HOC \( a_N \).

Furthermore, from Fig. 1 we should emphasize that the EKF-based estimator, of the amplitude \( A \), performs as well as the estimator (38) from SNR \( \geq 0 \) dB. In fact, since this last estimator is efficient from SNR \( \geq 0 \) dB...
(see Table 1) then the use of the EKF-based estimator, for estimating $A$, will not improve the performances.

The obtained results, which illustrate the high and very interesting performances of the EKF-based estimators, are due to the fact that the proposed method exploits implicitly the double of the initial number of snapshots. In fact, firstly the proposed step of initialization operates on $Ne$ snapshots and provides an initial state vector $x_0(0)$, nearest to the true values of $\{A, \theta, a_0, \ldots, a_N\}$ for low enough SNR, and secondly the EKF uses the same temporal snapshots to improve the performances of the estimation step (algorithm 2).

From all the figures, the obtained results show that the
EKF-based estimators exhibit performances which reach the CRB for low enough SNR. Thus, our method outperforms, in terms of statistical performances, JAHOCE proposed in [4] for only the estimation of the DOA $\theta$ and the HOC $a_N$.

Furthermore, from Fig. 1 we should emphasize that the EKF-based estimator, of the amplitude $A$, performs as well as the estimator (38) from SNR $\geq 0$ dB. In fact, since this last estimator is efficient from SNR $\geq 0$ dB (see Table 1) then the use of the EKF-based estimator, for estimating $A$, will not improve the performances.

The obtained results, which illustrate the high and very interesting performances of the EKF-based estimators, are due to the fact that the proposed method exploits implicitly the double of the initial number of snapshots. In fact, firstly the proposed step of initialization operates on $Ne$ snapshots and provides an initial state vector $x_p(0)$, nearest to the true values of $\{A, \theta, a_0, \ldots, a_N\}$ for low enough SNR, and secondly the EKF uses the same temporal snapshots to improve the performances of the estimation step (algorithm 2).

7 Conclusion

In this paper, we address the problem of estimating wideband polynomial phase signal (PPS) affected by additive noise. The principle of estimation is based on the introduction of an exact but nonlinear state model, of the PPS, which compels us to use the extended Kalman filter (EKF) which has never been used to solve such problems to the best of our knowledge. Furthermore, we propose a solution to the problem of initialization of the EKF since the initial conditions of the state model are assumed to be unknown. Under this initialization, which provides an initial state vector nearest to the true values of the unknown parameters, the numerical simulations show, for wideband QPS, that the proposed EKF-based estimators exhibit high performances and outperform in terms of statistical performances the proposed method in [4].

References


