A DFT Based Alternating Projection Algorithm for Parameter Estimation of Superimposed Complex Sinusoids

Gaurav Sharma and V. U. Reddy

Department of Electrical Communication Engineering
Indian Institute of Science
Bangalore 560 012, INDIA

Abstract

In this paper, we consider the alternating projection algorithm suggested by Ziskind and Wax [13] for parameter estimation of superimposed complex sinusoids (cisoids) embedded in noise, and show how the objective function being maximized at each step can be expressed as a non-linear function of the Fourier transforms of the observed data and discrete cisoids. We obtain elegant expressions for the objective functions for the case when the number of cisoids is two and develop a discrete Fourier transform (DFT) based algorithm for this case. The expressions for the case of three or more cisoids are quite cumbersome, but have a nice recursive structure that is illustrated by taking the case of three cisoids as an example. The details of the recursion for the general case are given in the appendix. We discuss some of the approximations and simplifications that lead to substantial reduction in computation. Simulation results are presented to show how the developed algorithm performs in comparison with the other techniques.

1 Introduction

The problem of estimating the parameters of superimposed complex sinusoids (cisoids) in noise from limited data has received considerable attention in the past few years. It is well known that the discrete Fourier transform (DFT) is an effective and efficient method of estimating the frequencies and the amplitudes of cisoids if there is only one cisoid, or if the frequencies are well separated with respect to the inverse of the observation interval. However, when there are multiple cisoids with closely spaced frequencies, DFT processing proves ineffective. In these cases, one has to resort to high resolution methods.

One class of high resolution methods is based on linear prediction formulation. Among the algorithms in this class, modified linear prediction approaches [10, 6] and the total least squares method [7, 8] have proved very effective. While the performance of these methods at high signal to noise ratio (SNR) is extremely good (often approaching the Cramér Rao bound (CRB)), they suffer from a drawback that their threshold SNR is reasonably large.

Another class of estimation algorithms is based on the non-linear least squares approach. If the noise is assumed to be white Gaussian, maximum likelihood (ML) estimation falls in this category. These algorithms exploit the information available about the signal to the maximum extent possible. While such algorithms perform well and have low threshold SNR, their computational complexity is too high to allow their use in practical applications. Hence, the emphasis in this class of algorithms has been towards simplifications and approximations that would make them computationally feasible. Several attempts have been made in this direction which have lead to iterative quadratic maximum likelihood (IQML) technique [5, 3], alternating projection (AP) algorithm [13, 2, 11] and constrained total least squares (CTLS) method [1].

In this paper, we consider the AP algorithm of Ziskind and Wax [13]. They formulated the dual of the cisoid parameter estimation problem arising in the direction of arrival estimation context as a non-linear least squares problem, and proposed the use of alternating maximization to convert the multi-dimensional maximization into a sequence of simpler one dimensional maximizations. In the DOA estimation scenario, with no constraints on the array response and geometry, they proposed a direct search for the maxima over the entire array manifold at each step. However, in the cisoid parameter estimation problem (with uniformly spaced samples), the structure in the problem can be exploited to implement the one dimensional search for maxima in an efficient manner. In this paper, we show how the objective function being maximized at each step can be expressed in terms of Fourier transforms of the data and of discrete cisoids, thereby allowing a DFT based implementation of the algorithm. Though the ideas of alternating maximization and projection matrix decomposition have been considered earlier, the idea of extended decomposition and the recognition that the resulting function can be expressed completely in terms of Fourier transforms are new. In addition, the proposed algorithm gives considerable computational saving over an AP algorithm involving a direct search over a grid of angular frequencies. Also, due to the structure of the algorithm, one needs to calculate the DFTs only once, and further manipulations may be done in the Fourier domain itself. The development of our algorithm has some similarity with reduced effort coarse search technique proposed by Van hamme [11]. But, our development goes beyond that of Van hamme as noted in Section 4.

The paper is organized as follows. In Section 2, we give the signal model under consideration and formulate the parameter estimation as a non-linear least squares problem. In Section 3, we give the development of the AP algorithm for the cisoid parameter estimation case. In Section 4, we develop the DFT based algorithm for two cisoids' case. In Section 5, we illustrate the structure of the algorithm for multiple cisoids by considering the three cisoids' case as an example. Computer simulation results comparing the performance of the proposed algorithm with other methods and with the CRB are presented in Section 6, and concluding remarks are given in Section 7.

2 Signal Model and Problem Statement

Consider a sequence of N uniformly spaced and noise corrupted samples from a signal consisting of M superimposed cisoids

$$y(n) = \sum_{i=1}^{M} s_i e^{j\,\omega_i\,n} + v(n), \quad n = 0, 1, \ldots, N-1 \tag{1}$$

where ω_i is the angular frequency of the i^{th} cisoid , s_i is its complex amplitude and $\{v(n)\}$ is a zero mean, stationary, complex valued random process with covariances

$$E[v(n)\,v^*(m)] = \sigma^2\,\delta_{n\,m} \tag{2}$$

$$E[v(n)\,v(m)] = 0 \tag{3}$$

where $\delta_{n\,m}$ denotes the Kronecker delta and superscript '*' denotes conjugation.

The problem to be solved is now stated as follows : assuming that the number of signals, M, is known, estimate the signal parameters $\{\omega_i\}_{i=1}^{M}$ and $\{s_i\}_{i=1}^{M}$ from the observed data.

Denoting the parameters to be estimated as

$$\boldsymbol{\omega} = [\,\omega_1,\,\omega_2,\,\ldots\,\omega_M\,]^T \tag{4}$$

$$\mathbf{s} = [\,s_1,\,s_2,\,\ldots\,s_M\,]^T \tag{5}$$

we formulate the estimation problem as

$$\hat{\boldsymbol{\omega}}\ ,\, \hat{\mathbf{s}} = \arg\min_{\boldsymbol{\omega}\,,\,\mathbf{s}} \sum_{n=0}^{N-1} \left| y(n) - \sum_{i=1}^{M} s_i e^{j\,\omega_i\,n} \right|^2 \tag{6}$$

Now, defining the vectors $\mathbf{y}$ and $\mathbf{a}(\omega)$ as

$$\mathbf{y} = [\, y(0),\, y(1),\, \ldots,\, y(N-1)\,]^T \tag{7}$$

$$\mathbf{a}(\omega) = [\, 1,\, e^{j\,\omega},\, e^{j\,2\,\omega}, \ldots\, e^{j\,(N-1)\,\omega}\,]^T \tag{8}$$

and the matrix $\mathbf{A}(\boldsymbol{\omega}\,)$ as

$$\mathbf{A}(\boldsymbol{\omega}\,) = [\, \mathbf{a}(\omega_1)\,,\, \mathbf{a}(\omega_2)\,,\, \ldots\,,\, \mathbf{a}(\omega_M)\,] \tag{9}$$

we can rewrite (6) as

$$\hat{\boldsymbol{\omega}}\,,\, \hat{\mathbf{s}} = \arg\min_{\boldsymbol{\omega}\,,\mathbf{s}} \|\, \mathbf{y} - \mathbf{A}(\boldsymbol{\omega}\,)\, \mathbf{s}\, \|^2 \tag{10}$$

where $\|\cdot\|$ denotes 2-norm. We may note here that this estimate corresponds to the ML estimate if the noise is Gaussian.

3 Alternating Projection (AP) Algorithm

The minimization in (10) is a multidimensional non-linear minimization in $3\,M$ real variables, and in its given form it is highly computation intensive. Reduction in the dimensionality of the problem can be achieved by noting that for any $\boldsymbol{\omega}$, the optimum value of $\mathbf{s}$ is given by

$$\hat{\mathbf{s}} = \mathbf{A}^{\#}(\boldsymbol{\omega}\,)\,\mathbf{y} \tag{11}$$

where superscript '#' denotes pseudoinverse. Substituting (11) in (10), we get

$$\hat{\boldsymbol{\omega}} = \arg\min_{\boldsymbol{\omega}} \|\, \mathbf{y} - \mathbf{P}_{\mathbf{A}(\boldsymbol{\omega}\,)}\, \mathbf{y}\, \|^2 \tag{12}$$

$$\hat{\mathbf{s}} = \left(\mathbf{A}^{+}(\hat{\boldsymbol{\omega}}\,)\,\mathbf{A}(\hat{\boldsymbol{\omega}}\,)\right)^{-1} \mathbf{A}^{+}(\hat{\boldsymbol{\omega}}\,)\,\mathbf{y} \tag{13}$$

where superscript '+' denotes Hermitian transpose and $\mathbf{P}_{\mathbf{A}(\boldsymbol{\omega}\,)}$ is the projection matrix which projects a vector onto the column space of $\mathbf{A}(\boldsymbol{\omega}\,)$. $\mathbf{P}_{\mathbf{A}(\boldsymbol{\omega}\,)}$ is given by

$$\mathbf{P}_{\mathbf{A}(\boldsymbol{\omega}\,)} = \mathbf{A}(\boldsymbol{\omega}\,)\,(\mathbf{A}^{+}(\boldsymbol{\omega}\,)\mathbf{A}(\boldsymbol{\omega}\,))^{-1}\mathbf{A}^{+}(\boldsymbol{\omega}\,) \tag{14}$$

Equation (12) may alternatively be written as

$$\begin{aligned}\hat{\boldsymbol{\omega}} &= \arg \max_{\boldsymbol{\omega}} \ \| \mathbf{P}_{\mathbf{A}(\boldsymbol{\omega})} \, \mathbf{y} \|^2 \\ &= \arg \max_{\boldsymbol{\omega}} \ \mathbf{y}^{+} \mathbf{P}_{\mathbf{A}(\boldsymbol{\omega})} \, \mathbf{y} \end{aligned} \tag{15}$$

This is a M dimensional nonlinear maximization and is still computationally expensive. Using the alternating maximization technique, Ziskind and Wax [13] transformed it into a sequence of simpler one-dimensional problems. The technique is iterative; at every step a maximization is performed with respect to a single parameter while all other parameters are held fixed. That is, the estimate of ω_i at $(k+1)^{th}$ iteration is obtained by solving the following one-dimensional maximization problem

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\omega} \ \mathbf{y}^{+} \mathbf{P}_{[\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)}), \mathbf{a}(\omega)]} \, \mathbf{y} \tag{16}$$

where $\tilde{\boldsymbol{\omega}}_i^{(k)}$ denotes the $(M-1) \times 1$ vector of the pre-estimated parameters

$$\tilde{\boldsymbol{\omega}}_i^{(k)} = [\, \hat{\omega}_1^{(k+1)} , \ \hat{\omega}_2^{(k+1)} , \ldots , \hat{\omega}_{i-1}^{(k+1)} , \ \hat{\omega}_{i+1}^{(k)} , \ldots , \hat{\omega}_M^{(k)} \,]^T \tag{17}$$

For initializing the algorithm we begin by solving the problem for a single cisoid,

$$\hat{\omega}_1^{(0)} = \arg \max_{\omega} \ \mathbf{y}^{+} \mathbf{P}_{\mathbf{a}(\omega)} \, \mathbf{y} \tag{18}$$

Next, we solve for $\hat{\omega}_2^{(0)}$ with the first cisoid angular frequency fixed at $\hat{\omega}_1^{(0)}$,

$$\hat{\omega}_2^{(0)} = \arg \max_{\omega} \ \mathbf{y}^{+} \mathbf{P}_{[\mathbf{a}(\hat{\omega}_1^{(0)}), \mathbf{a}(\omega)]} \, \mathbf{y} \tag{19}$$

Continuing in this fashion at i^{th} initialization step, we determine the initial estimate of the i^{th} angular frequency, $\hat{\omega}_i{}^{(0)}$, assuming that there are only i cisoids and $(i-1)$ of these are located at their pre-estimated frequencies $\hat{\omega}_1^{(0)}, \hat{\omega}_2^{(0)}, \ldots, \hat{\omega}_{i-1}^{(0)}$. The procedure is continued till all the initial values are determined.

The maximization in (16) still involves considerable computation due to the matrix inversion and matrix-matrix multiplication required at each step. We now use a basic property of projection matrices (known as the projection-matrix update formula [13]) to reduce the computation further.

3.1 Projection Matrix Decomposition

Let $\mathbf{B}$ and $\mathbf{C}$ be two arbitrary matrices with the same number of rows, and let $\mathbf{P}_{[\mathbf{B},\mathbf{C}]}$ denote the projection-matrix onto the column space of the augmented matrix $[\mathbf{B},\mathbf{C}]$. Then from the decomposition given by Ziskind and Wax [13] we have

$$\mathbf{P}_{[\mathbf{B},\mathbf{C}]} = \mathbf{P}_{\mathbf{B}} + \mathbf{P}_{\mathbf{C}_{\mathbf{B}}} \tag{20}$$

where $\mathbf{C}_{\mathbf{B}}$ denotes the residual of the columns of $\mathbf{C}$ when projected onto the column space of $\mathbf{B}$, and is given by

$$\mathbf{C}_{\mathbf{B}} = (\mathbf{I} - \mathbf{P}_{\mathbf{B}})\mathbf{C} \tag{21}$$

Applying (20) to the projection matrix in (16), we get

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\omega} \left\{ \mathbf{y}^+ \left(\mathbf{P}_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})} + \mathbf{P}_{\mathbf{a}(\omega)_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}} \right) \mathbf{y} \right\} \tag{22}$$

Since the first term does not depend on ω, (22) reduces to

$$\hat{\omega}_i^{(k+1)} = \arg \max_{\omega} \left\{ \mathbf{y}^+ \mathbf{P}_{\mathbf{a}(\omega)_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}} \mathbf{y} \right\} \tag{23}$$

where

$$\mathbf{a}(\omega)_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})} = \left(I - \mathbf{P}_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})} \right) \mathbf{a}(\omega) \tag{24}$$

4 Two Cisoids' Case

This is the simplest case with which both the issues of resolution and estimation accuracy can be addressed. For this case,

$$\mathbf{A}(\tilde{\boldsymbol{\omega}}\,_i^{(k)}) = \begin{cases} \mathbf{a}(\hat{\omega}_2^{(k)}) & i = 1 \\ \mathbf{a}(\hat{\omega}_1^{(k+1)}) & i = 2 \end{cases} \tag{25}$$

For any non-zero vector $\mathbf{u}$

$$\mathbf{P_u} = \frac{\mathbf{u}\,\mathbf{u}^+}{(\mathbf{u}^+\,\mathbf{u})} \tag{26}$$

and in view of (8), we have

$$\mathbf{a}^+(\omega)\,\mathbf{y} \;=\; \sum_{n=0}^{N-1} y(n)\,e^{-j\omega n} \;=\; Y(\omega) \tag{27}$$

$$\mathbf{a}^+(\omega)\,\mathbf{a}(\xi) \;=\; U(\omega - \xi) \tag{28}$$

where $Y(\omega)$ denotes the Fourier transform of $\{y(n)\}$ and

$$U(\omega) \;=\; \sum_{n=0}^{N-1} e^{-j\,\omega\,n} \;=\; \begin{cases} \frac{e^{j\,(\frac{N-1}{2})\,\omega}\;sin(\frac{N\,\omega}{2})}{sin(\frac{\omega}{2})} & \omega \neq 0 \\ N & \omega = 0 \end{cases} \tag{29}$$

Note that $U(\omega)$ is the Fourier transform of a sequence of N ones, i.e., a discrete cisoid sequence of unit amplitude and zero angular frequency. We now develop the proposed algorithm for the case of two cisoids.

Using the relations (26) to (28) we can show that

$$\mathbf{y}^+\,\mathbf{P}_{\mathbf{a}(\omega)\mathbf{a}(\xi)}\,\mathbf{y} = \frac{\left|Y(\omega) - \frac{Y(\xi)\,U(\omega-\xi)}{N}\right|^2}{\left(N - \frac{|U(\omega-\xi)|^2}{N}\right)} \tag{30}$$

Thus, the AP algorithm for the case of two cisoids is as follows:

(1) INITIALIZATION :

$$k \;=\; 0 \tag{31}$$

$$\hat{\omega}_1^{(0)} \;=\; \arg\,\max_{\omega} \frac{|\,Y(\omega)\,|^2}{N} \tag{32}$$

(2) ITERATIVE STEP :

$$\hat{\omega_2}^{(k)} = \arg \max_{\omega} \left\{ \frac{\left| Y(\omega) - \frac{Y(\hat{\omega}_1^{(k)})}{N} U(\omega - \hat{\omega}_1^{(k)}) \right|^2}{\left(N - \frac{\left| U(\omega - \hat{\omega}_1^{(k)}) \right|^2}{N} \right)} \right\} \tag{33}$$

$$\hat{\omega_1}^{(k+1)} = \arg \max_{\omega} \left\{ \frac{\left| Y(\omega) - \frac{Y(\hat{\omega}_2^{(k)})}{N} U(\omega - \hat{\omega}_2^{(k)}) \right|^2}{\left(N - \frac{\left| U(\omega - \hat{\omega}_2^{(k)}) \right|^2}{N} \right)} \right\} \tag{34}$$

$$k = k + 1 \tag{35}$$

(3) CONVERGENCE CHECK :

If $\| \, (\hat{\boldsymbol{\omega}}^{(k)} - \hat{\boldsymbol{\omega}}^{(k-1)}) \, \|_\infty < \varepsilon$ *stop* ; *else* *go to step* (2)

where $\| \cdot \|_\infty$ denotes ∞-norm and ε is some suitably chosen small convergence parameter.

We note that $Y(\omega)$ is the Fourier transform of the given data sequence and $U(\omega - \xi)$ is the Fourier transform of a unit amplitude discrete cisoid sequence of length N having frequency ξ. Thus, the term $| \ldots |$ in the numerator in step (2) (*cf.* (33)) has an intuitively appealing physical interpretation as the Fourier transform of a modified signal obtained by subtracting a discrete cisoid of frequency $\hat{\omega}_1^{(k)}$ and complex amplitude $Y(\hat{\omega}_1^{(k)})/N$ from the original sequence. However, the denominator is quite non-intuitive and it arises due to the fact that when two cisoid frequencies are close they become coupled. Hence, even if $\hat{\omega}_1^{(k)}$ is the true frequency of the first cisoid, $Y(\hat{\omega}_1^{(k)})/N$ is not its true amplitude.

We may note here that (30) is similar to Eqn. (8) in Van hamme [11]. However, our development goes beyond the treatment of Van hamme and expresses the objective function completely in terms of Fourier transforms. This can be seen more clearly for the three cisoids case which we shall treat in the next section.

4.1 DFT Based Implementation

In order to implement the minimization in a computationally efficient manner, we discretize the search. So as to take advantage of fast DFT algorithms, we carry out the search for maxima at frequencies given by $(2\,\pi\,q)/L, \quad q = 0, 1 \ldots L-1$. L should be large so that the error introduced due to discretization is small compared to the root mean squared estimation error. Now, instead of $\hat{\omega}_1^{(k)}$ and $\hat{\omega}_2^{(k)}$ we estimate $q_1^{(k)}$ and $q_2^{(k)}$ where

$$\hat{\omega}_1^{(k)} = \frac{2\,\pi\,q_1^{(k)}}{L} \tag{36}$$

$$\hat{\omega}_2^{(k)} = \frac{2\,\pi\,q_2^{(k)}}{L} \tag{37}$$

Denote

$$Y\left(\frac{2\,\pi\,q}{L}\right) = \overline{Y}(q) \tag{38}$$

$$U\left(\frac{2\,\pi\,q}{L}\right) = \overline{U}(q) \tag{39}$$

The DFT based AP algorithm is given as follows :

(1) INITIALIZATION :

$$k = 0 \tag{40}$$

$$q_1^{(0)} = \arg\,\max_q \;|\,\tilde{Y}(q)\,|^2 \tag{41}$$

(2) ITERATIVE STEP :

$$q_2^{(k)} = \arg\;\max_q \left\{ \frac{\left|\tilde{Y}(q) - \tilde{Y}(q_1^{(k)})\,\tilde{U}(q - q_1^{(k)})\right|^2}{\left(1 - \left|\tilde{U}(q - q_1^{(k)})\right|^2\right)} \right\} \tag{42}$$

$$q_1^{(k+1)} = \arg\;\max_q \left\{ \frac{\left|\tilde{Y}(q) - \tilde{Y}(q_2^{(k)})\,\tilde{U}(q - q_2^{(k)})\right|^2}{\left(1 - \left|\tilde{U}(q - q_2^{(k)})\right|^2\right)} \right\} \tag{43}$$

$$k = k + 1 \tag{44}$$

(3) CONVERGENCE CHECK :

If $q_1^{(k)} = q_1^{(k-1)}$ *stop* ; *else* *go to step* (2)

In (41) to (43),

$$\tilde{Y}(q) = \frac{\overline{Y}(q)}{N} = \text{DFT of sequence } \left\{\frac{y(n)}{N}\right\}_{n=0}^{N-1} \tag{45}$$

$$\tilde{U}(q) = \frac{\overline{U}(q)}{N} = \text{DFT of sequence } \underbrace{\left\{\frac{1}{N}, \frac{1}{N}, \ldots \frac{1}{N}\right\}}_{N} \tag{46}$$

Considerable computational savings can be achieved here. Firstly, $\{\tilde{Y}(q)\}_{q=0}^{L-1}$ corresponds to the L point DFT of the N point sequence $\left\{\frac{y(n)}{N}\right\}_{n=0}^{N-1}$ appended with $L-N$ zeros, which may be evaluated at the start of the algorithm and stored. Similarly, $\{\tilde{U}(q)\}_{q=0}^{L-1}$ corresponds to the L point DFT of $\underbrace{\left\{\frac{1}{N}, \frac{1}{N}, \ldots \frac{1}{N}\right\}}_{N}$ appended with $L-N$ zeros, and can be computed and stored beforehand. Further, since typically N is very small in comparison with L, we can use pruned fast Fourier transform (FFT) algorithms for the computation. $\{\tilde{U}(q-l)\}_{q=0}^{L-1}$ simply corresponds to $\{\tilde{U}(q)\}_{q=0}^{L-1}$ with a right circular shift of l positions. The plot of $\mid \tilde{U}(q) \mid$ in dB as a function of q for $N = 25$ and $L = 2048$ is given in Fig 1. The graph is identical to the magnitude of the Fourier transform of a rectangular window. In general, the pattern has a main lobe of width $(2\,L/N)$ at $q = 0$ (taking into account the circular fold-over in the DFT) and sidelobes of width (L/N). The peak amplitude of the main lobe is 0 dB while that of the first sidelobe is approximately -13 dB, second and higher order sidelobes are less than -17 dB. Hence, to a reasonable accuracy, $\tilde{U}(q)$ may be approximated by a $(4\,L/N)$ length sequence comprising of the main lobe and one sidelobe on either side. This approximation is always valid in the denominators of (42) and (43). This can be seen from Fig. 2 where $1/(1 - |\tilde{U}(q)|^2)$ has been plotted in dB as a function of q. The approximation is valid in the numerator too if the difference in amplitudes of the two cisoids is small. Alternatively, different approximations involving more or less sidelobes may be used in the numerator depending on the approximate knowledge of the relative amplitudes. Using these approximations, considerable saving in computation can be achieved with negligible degradation in the performance.

5 Extension of the Algorithm to Multiple Cisoids

In this section, we extend the AP algorithm to the case of multiple cisoids. The expressions for the case of two cisoids, developed in the previous section, are quite elegant. However, the expressions for three or more cisoids tend to get quite lengthy, though they have a simple recursive structure allowing one to program the general case.

5.1 Extended Projection Matrix Decomposition

The idea of projection matrix decomposition, given in Section 3, can be applied repeatedly to the projection matrix $\mathbf{P}_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}$, removing a single column at each stage till only one column is left.

In our case, $\left\{\mathbf{y}^+ \left(\mathbf{P}_{\mathbf{a}(\omega)_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}}\right) \mathbf{y}\right\}$ in (23) can be expressed as

$$\mathbf{y}^+ \left(\mathbf{P}_{\mathbf{a}(\omega)_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}}\right) \mathbf{y} = \frac{\left|\mathbf{a}(\omega)^+_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})} \mathbf{y}\right|^2}{\|\ \mathbf{a}(\omega)_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}\ \|^2} \tag{47}$$

Simplification of this expression using repeated decomposition of $\mathbf{P}_{\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})}$ can be achieved much more readily than the direct expansion of the projection matrix as all the columns of $\mathbf{A}(\tilde{\boldsymbol{\omega}}_i^{(k)})$ have the same constant norm N, and the quadratic form makes simplification easy. We illustrate the general pattern of (47) considering the case of three cisoids as an example. The general case of multiple cisoids is considered in the appendix.

5.2 Three Cisoids' Case

Repeated application of the decomposition of (20) and (21) to the expression in (47) yields the AP algorithm for the case of three cisoids. Once again, as in Section 4, we carry out the search for maxima at frequencies given by $(2\ \pi\ q)/L, \quad q = 0, 1 \ldots L-1$, (instead of $\hat{\omega}_1^{(k)}$, $\hat{\omega}_2^{(k)}$ and $\hat{\omega}_3^{(k)}$) and estimate $q_1^{(k)}$, $q_2^{(k)}$ and $q_3^{(k)}$ where $\hat{\omega}_1^{(k)} = \frac{2\ \pi\ q_1^{(k)}}{L}$, $\hat{\omega}_2^{(k)} = \frac{2\ \pi\ q_2^{(k)}}{L}$ and $\hat{\omega}_3^{(k)} = \frac{2\ \pi\ q_3^{(k)}}{L}$.

Defining $\tilde{Y}$, $\tilde{U}$ as in Section 4, we can now express the DFT based AP Algorithm for three cisoids as follows :

(1) INITIALIZATION :

$$k = 0 \tag{48}$$

$$q_1^{(0)} = \arg \max_q \mid \tilde{Y}(q) \mid^2 \tag{49}$$

$$q_2^{(0)} = \arg \max_q \left\{ \frac{\left| \tilde{Y}(q) - \tilde{Y}(q_1^{(0)})\, \tilde{U}(q - q_1^{(0)}) \right|^2}{\left(1 - \left| \tilde{U}(q - q_1^{(0)}) \right|^2\right)} \right\} \tag{50}$$

(2) ITERATIVE STEP :

$$
\begin{aligned}
q_3^{(k)} &= \arg \max_q \left\{ \frac{\left| \tilde{Y}(q) - \tilde{Y}(q_1^{(k)})\, \tilde{U}(q - q_1^{(k)}) - \frac{\tilde{Y}(q_2^{(k)}) - \tilde{Y}(q_1^{(k)})\, \tilde{U}(q_2^{(k)} - q_1^{(k)})}{\left(1 - \left|\tilde{U}(q_2^{(k)} - q_1^{(k)})\right|^2\right)} \left(\tilde{U}(q - q_2^{(k)}) - \tilde{U}(q_1^{(k)} - q_2^{(k)}) \cdots \right.\right.}{\left(1 - \left|\tilde{U}(q - q_1^{(k)})\right|^2 - \frac{\left|\tilde{U}(q - q_2^{(k)}) - \tilde{U}(q_1^{(k)} - q_2^{(k)})\, \tilde{U}(q - q_1^{(k)})\right|^2}{\left(1 - \left|\tilde{U}(q_2^{(k)} - q_1^{(k)})\right|^2\right)}\right)} \right. \\
q_1^{(k+1)} &= \arg \max_q \left\{ \frac{\left| \tilde{Y}(q) - \tilde{Y}(q_2^{(k)})\, \tilde{U}(q - q_2^{(k)}) - \frac{\tilde{Y}(q_3^{(k)}) - \tilde{Y}(q_2^{(k)})\, \tilde{U}(q_3^{(k)} - q_2^{(k)})}{\left(1 - \left|\tilde{U}(q_3^{(k)} - q_2^{(k)})\right|^2\right)} \left(\tilde{U}(q - q_3^{(k)}) - \tilde{U}(q_2^{(k)} - q_3^{(k)}) \cdots \right.\right.}{\left(1 - \left|\tilde{U}(q - q_2^{(k)})\right|^2 - \frac{\left|\tilde{U}(q - q_3^{(k)}) - \tilde{U}(q_2^{(k)} - q_3^{(k)})\, \tilde{U}(q - q_2^{(k)})\right|^2}{\left(1 - \left|\tilde{U}(q_3^{(k)} - q_2^{(k)})\right|^2\right)}\right)} \right. \\
q_2^{(k+1)} &= \arg \max_q \left\{ \frac{\left| \tilde{Y}(q) - \tilde{Y}(q_3^{(k)})\, \tilde{U}(q - q_3^{(k)}) - \frac{\tilde{Y}(q_1^{(k+1)}) - \tilde{Y}(q_3^{(k)})\, \tilde{U}(q_1^{(k)} - q_3^{(k)})}{\left(1 - \left|\tilde{U}(q_1^{(k+1)} - q_3^{(k)})\right|^2\right)} \left(\tilde{U}(q - q_1^{(k+1)}) - \tilde{U}(q_3^{(k)} - \cdots \right.\right.}{\left(1 - \left|\tilde{U}(q - q_3^{(k)})\right|^2 - \frac{\left|\tilde{U}(q - q_1^{(k+1)}) - \tilde{U}(q_3^{(k)} - q_1^{(k+1)})\, \tilde{U}(q - q_3^{(k)})\right|^2}{\left(1 - \left|\tilde{U}(q_1^{(k+1)} - q_3^{(k)})\right|^2\right)}\right)} \right. \\
k &= k + 1
\end{aligned}
$$

(3) CONVERGENCE CHECK :

If $q_1^{(k)} = q_1^{(k-1)}$ *and* $q_2^{(k)} = q_2^{(k-1)}$ *stop* ; *else* *go to step* (2)

5.3 Computational Simplifications

The computational simplifications mentioned in Section 4 are still valid for the case of multiple cisoids. Further, several other simplifications can also be made. In view of the limited region in which $\tilde{U}$ is significant, the region in which the product terms involving two or more terms in $\tilde{U}$ are significant is still smaller. We also note that if two groups of angular frequencies $\omega_1, \omega_2, \ldots \omega_k$ and $\xi_1, \xi_2, \ldots \xi_r$ are widely separated (as compared to $2\,\pi/N$), then one can write

$$\mathbf{P}_{[\mathbf{a}(\omega_1),\mathbf{a}(\omega_2),\ldots\mathbf{a}(\omega_k),\mathbf{a}(\xi_1),\mathbf{a}(\xi_2),\ldots,\mathbf{a}(\xi_r)]} \simeq \mathbf{P}_{[\mathbf{a}(\omega_1),\ldots,\mathbf{a}(\omega_k)]} + \mathbf{P}_{[\mathbf{a}(\xi_1),\ldots,\mathbf{a}(\xi_r)]} \tag{55}$$

This can be used while decomposing the projection matrix so as to reduce the complexity of the terms involved. However, while the above simplification is useful, we still have ω as a variable and hence $\mathbf{a}(\omega)$ cannot be decoupled from the other columns of $\mathbf{A}(\tilde{\boldsymbol{\omega}}\,_i^{(k)})$ (see (23) and (24)). This problem can also be circumvented and further reduction in computation can be achieved by searching for local maxima in the neighbourhood of the previous estimate rather than searching over the entire range (at each step). In such case, the original equation for the estimate of the i^{th} frequency at the $k+1^{th}$ iteration can be written as

$$\hat{\omega}_i^{(k+1)} = \arg\max_{\omega\in\Omega}\ \mathbf{y}^{+}\,\mathbf{P}_{[\mathbf{A}(\tilde{\boldsymbol{\omega}}\,_i^{(k)}),\,\mathbf{a}(\omega)]}\,\mathbf{y} \tag{56}$$

where $\tilde{\boldsymbol{\omega}}\,_i^{(k)}$ denotes the $(M-1)\times 1$ vector of the pre-estimated parameters

$$\tilde{\boldsymbol{\omega}}\,_i^{(k)} = [\hat{\omega}_1^{(k+1)}\,,\ \hat{\omega}_2^{(k+1)}\,,\ldots,\hat{\omega}_{i-1}^{(k+1)}\,,\ \hat{\omega}_{i+1}^{(k)}\,,\ldots,\hat{\omega}_M^{(k)}]^T \tag{57}$$

and Ω is a set including $\hat{\omega}_i^{(k)}$ and the neighbourhood around it in which the search is to be conducted. Now using the decomposition given in (55), we can drop from $\mathbf{A}(\tilde{\boldsymbol{\omega}}\,_i^{(k)})$ the columns corresponding to angular frequencies far away from Ω.

6 Simulation Results

In this section we present simulation results to show how the proposed algorithm performs in relation to the existing algorithms [10, 6, 7, 3] and the theoretical best performance predicted by the Cramér Rao (CR) bound.

The signal model considered was

$$y(n) = e^{j(2\pi f_1 n + \phi_1)} + e^{j(2\pi f_2 n + \phi_2)} + v(n), \qquad n = 0, 1, \ldots, 24$$

with $f_1 = 0.5, f_2 = 0.52$ (normalized frequencies), $\phi_1 = 0$ and $\phi_2 = \pi/4$. $\{v(n)\}$ was a sequence of independent and identically distributed zero mean complex Gaussian random variables with uncorrelated real and imaginary parts, each with a variance of $\sigma^2/2$. The value of σ was chosen to give the desired SNR, defined as $10\, log_{10} \frac{1}{\sigma^2}$. We may mention here that while the parameter values are the same as in [10, 6, 7, 3], we have the index n running from 0 to 24 while [10, 6, 7, 3] use $n = 1, 2 \ldots 25$.

For the DFT based AP algorithm two values of L (2048 and 4096) were chosen. The choice of powers of 2 enables one to use fast radix 2 FFT algorithms. For MFBLP and TLS, the predictor order L was chosen as 18. This corresponds to $(3\,N)/4$ which was shown (experimentally) to be the near optimum value by Tufts and Kumaresan [10]. In the IQML method, the convergence test parameter was fixed at 5×10^{-6}.

The mean square error (MSE) in the frequency estimates was evaluated from 50 Monte-Carlo runs. This was repeated for different values of SNR and the results are plotted in Fig. 3 for the case $L = 2048$ and in Fig. 4 for the case $L = 4096$. In the same figures, we also overlaid the MSE values obtained with the IQML method of Bresler and Macovski, with MFBLP and TLS methods for the predictor order 18. All the algorithms were simulated using the same 50 noise realizations with appropriate scaling for different SNR values. The figures also contain the CR bounds.

The plots of Fig. 3 show that among the methods for which the results are presented, the DFT based AP algorithm yields lowest threshold SNR; 3 dB less than that of IQML and 6 dB less than that of MFBLP and TLS in the case of the estimate of f_1 (0.5), while the corresponding values are 2 dB and 7 dB in the case of f_2 (0.52).

Above the threshold SNR, the performance of the DFT based AP algorithm (DFTAP) is superior to that of MFBLP and TLS methods by 1 to 2 dB and is similar to that of IQML. Note that the TLS and MFBLP perform similarly which is consistent with the results given in [8] (In fact, the two curves cannot be resolved in Figs. 3 and 4). On average the DFT based AP algorithm required 3-4 iterations for it to converge at high SNR's and 4-5 iterations at SNR's close to the threshold SNR.

The plot corresponding to the DFTAP shows some fluctuations which are expected because the effect of discretization may not be averaged out over the 50 Monte-Carlo runs. In addition, at SNR greater than 30 dB, DFTAP for $L = 2048$ begins to show saturation as the value of the error introduced by discretization inherent in the algorithm is no longer negligible in comparison to the MSE. This explanation is also borne out by the curves for $L = 4096$ in Fig. 4, where we see that the saturation sets in much later as we have a finer discretization. Also, as expected theoretically, the effects of saturation become apparent approximately at the point where $1/L$ equals the root mean square value corresponding to the CRB. To overcome this saturation effect, one could (as a final step) have a finer grid of Y around $\hat{\omega}_1^{(k)}$, $\hat{\omega}_2^{(k)}$, $\ldots \hat{\omega}_M^{(k)}$ (the final DFTAP estimates) and of U around the pairwise differences of these angular frequencies. This requires direct evaluation of the Fourier transforms at the required points. But, it is still computationally attractive because N is quite small, and hence the computational effort in this evaluation is much less than that required for increasing the value of L *per se*. Finally, the MSE values better than the CR bound, observed for DFTAP in the neighbourhood of the threshold SNR, may have resulted because of small number of Monte-Carlo runs or due to slight bias in the estimates due to the discrete nature of the algorithm.

7 Conclusions

In this paper, we addressed the problem of parameter estimation of superimposed cisoids in noise using the alternating projection (AP) algorithm. Starting with the AP algorithm it was shown how the objective function can be expressed in terms of Fourier transforms. A DFT based implementation of the AP algorithm was then presented for the two cisoids' case. Several computational simplifications were suggested for this case. The algorithm was then extended to the case of multiple cisoids taking the case of three cisoids as an illustrative example. Additional computational simplifications possible for the multiple cisoids case were discussed. Simulation results were presented comparing the performance of the developed algorithm with the MFBLP method, total least squares approach and the IQML algorithm. The algorithm was seen to have a lower threshold SNR and the performance above the threshold was similar to that of the other methods.

Appendix

In this appendix we demonstrate the recursive structure of the algorithm for the general case of multiple cisoids. For this purpose we shall consider the initialization step of the AP algorithm. Note that the objective functions for the maximizations in the iterative step can be readily obtained by a cyclic permutation of the frequencies in the objective function for the final initialization step.

The initial estimate for the first frequency is given by

$$\hat{q}_1 = \arg\max_q \frac{|\ V_1(q)\ |^2}{D_1(q)}$$

where

$$\begin{aligned} V_1(q) &= \tilde{Y}(q) \\ D_1(q) &= 1 \\ W_1(q) &= \tilde{U}(q - \hat{q}_1) \end{aligned}$$

and the initial estimates for the other frequencies are given by the following recursive algorithm:

for $n = 1$ to $M - 1$ do

$$\begin{aligned} V_{n+1}(q) &= V_n(q) - \frac{V_n(\hat{q}_n)}{D_n(\hat{q}_n)} W_n(q) \\ D_{n+1}(q) &= D_n(q) - \frac{|\ W_n(q)\ |^2}{D_n(\hat{q}_n)} \\ \hat{q}_{n+1} &= \arg\max_q \frac{|\ V_{n+1}(q)\ |^2}{D_{n+1}(q)} \\ W_{n+1}(q) &= \tilde{U}(q - \hat{q}_{n+1}) - \frac{W_n^*(\hat{q}_{n+1})}{D_n(\hat{q}_n)} W_n(q) \end{aligned}$$

end { for loop }.

The iterative step is now obtained readily by a cyclic permutation of the estimated frequencies in the objective function for the initialization of $\hat{q}_M$. From the recursion it can also be seen that the algorithm requires $(3\,M^2 - 2\,M)\,L$ multiplications/divisions and $3\,M\,(M - 1)\,L$

additions/subtractions per iteration of the AP algorithm (where M is the number of cisoids and L is the length of the DFT used).

References

[1] T. J. Abatzoglou, Jerry M. Mendel and G. A. Harada, “ The Constrained Total Least Squares Technique and its Application to Harmonic Superresolution”, *IEEE Trans. on Signal Processing*, Vol. 39, No. 5, May 1991, pp. 1070-1087.

[2] L.S. Biradar and V.U. Reddy, “A Newton-based Ziskind-Wax AP algorithm for parameter estimation exponential signals in noise ”, *IEEE Trans. on SP*, Vol. 41, No. 3, March 1993, pp. 1435-1438. .

[3] Y. Bresler and A. Macovski, “Exact maximum likelihood parameter estimation of superimposed exponential signals in noise”, *IEEE Trans. on ASSP*, Vol. 34, No. 5, Oct. 1986, pp. 1081-1089.

[4] G.H. Golub and V. Pereyra, “The differentiation of pseudo-inverses and nonlinear least-squares problems whose variables separate”, *SIAM J. Numer. Anal.*, Vol. 10, No. 2, April 1973, pp. 413-432.

[5] R. Kumaresan, L.L. Scharf and A.K. Shaw, “An algorithm for pole-zero modelling and spectral analysis”, *IEEE Trans. on ASSP*, Vol. 34, No. 3, June 1986, pp. 637-640.

[6] R. Kumaresan and D.W. Tufts, “Estimating the parameters of exponentially damped sinusoids and pole-zero modelling in noise”, *IEEE Trans. on ASSP*, Vol. 30, No. 6, Dec. 1982, pp. 833-840.

[7] Md. A. Rahman and K. Yu, “Total least squares approach for frequency estimation using linear prediction”, *IEEE Trans on ASSP*, Vol. 35, No. 10, Oct. 1987, pp. 1440-1454.

[8] V.U. Reddy and B. Srinivas, “Modified forward-backward linear prediction: Improvements and criterion for detection of the number of sinusoids”, *J. IETE (India)*, Vol. 35, No. 2, March-April 1989, pp. 85-92 (Invited Paper).

[9] D. Starer and A. Nehorai, “ Maximum Likelihood Estimation of Exponential Signals in Noise using a Newton Algorithm”, *Fourth ASSP Workshop on Spectral Estimation and Modelling,*Minneapolis Aug. 1988.

[10] D.W. Tufts and R. Kumaresan, “Estimation of frequencies of multiple sinusoids : Making linear prediction perform like maximum likelihood”, *Proc. IEEE*, Vol. 70, Sept. 1982, pp. 975-989.

[11] H. Van hamme, “Maximum likelihood estimation of superimposed complex sinusoids in white Gaussian noise by Reduced Effort Coarse Search (RECS)”, *IEEE Trans. on SP*, Vol. 39, No. 2, Feb. 1991, pp. 536-538.

[12] M. Wax, Detection and estimation of superimposed signals, *Ph.D dissertation*, Stanford Univ., Stanford, CA, 1985.

[13] I. Ziskind and M. Wax, “Maximum likelihood localization of multiple sources by alternating projection”, *IEEE Trans. on ASSP*, Vol. 36, No. 10, Oct. 1988, pp. 1553-1560.