A DOA ESTIMATOR BASED ON LINEAR PREDICTION AND TOTAL LEAST SQUARES

Amauri Lopes, Ivanil S. Bonatti, Pedro L. D. Peres, Ricardo F. Colares and Carlos A. Alves

Abstract - We propose an estimator for the direction of arrival (DOA) of plane waves incident on a linear equally spaced array of sensors. The estimator uses a forward - backward linear prediction filter optimized by means of the total least squares criterion subject to constraints. The DOA angles are estimated by the zeros of the optimum filter. The proposed procedure is compared with two maximum likelihood methods presented in the literature.

Keywords: Array Signal Processing, Direction of Arrival, Linear Prediction, Total Least Squares.

Resumo - Propomos um estimador para a direção de chegada (DOA) de ondas planas incidentes em um arranjo de sensores linear e com elementos igualmente espaçados. O estimador usa um filtro de predição linear "forward-backward" otimizado através do critério dos mínimos quadrados totais sujeito a restrições. Os ângulos DOA são estimados através dos zeros do filtro otimizado. O procedimento proposto é comparado com dois métodos de máxima verossimilhança presentes na literatura.

Palavras-chave: Processamento de Sinais de Sensores, Direção de Chegada, Predição Linear, Mínimos Quadrados Totais.

1. INTRODUCTION

Array signal processing is a subarea of statistical signal processing and its main purpose is the estimation of signal parameters by fusing temporal and spatial (space-time processing) information obtained via sampling a wavefield with a set of sensors. The sensors are conveniently arranged in space and the wavefield is assumed to be generated by a finite number of emitters.

One of the applications of array signal processing is the estimation of the direction of arrival (DOA) of the waves generated by the emitters. The DOA problem and array signal processing are present in mobile communication, radar, sonar, seismology, radio astronomy and industry applications [1].

This paper deals with DOA estimation of plane waves incident on a linear equally spaced array of sensors.

Amauri Lopes, Ivanil S. Bonatti and Pedro L. D. Peres are with the University of Campinas. E-mails: amauri@decom.fee.unicamp.br, (ivanil@, peres@)dt.fee.unicamp.br. Ricardo F. Colares is with the Universidade de Fortaleza (UNIFOR). E-mail: ricardocolares@unifor.br. Carlos A. Alves is with the São Paulo State University/Ilha Solteira. E-mail: caa@dee.feis.unesp.br.

When the DOA angles are well separated in space, discrete Fourier transform (DFT) is efficient and effective for DOA estimation [2]. However, if the angles are more closely spaced than the reciprocal of the aperture in space, the DFT is ineffective and a high resolution technique is required. Direct maximum likelihood (ML) estimation is effective but being a non-linear least squares fitting problem, requires a multidimensional search procedure to estimate the angles. Recently, ML-based methods were proposed [3], [4], where a reparameterization of the direct ML problem is employed to circumvent the multidimensional search. This reparameterization leads to an optimization problem, which is solved by means of an iterative algorithm. MODE [5] and MODEX [6] are meaningful examples of this class of estimators.

As an alternative to the reparameterised ML estimators, linear prediction leads to an estimator that can be used if the signal-to-noise ratio (SNR) is sufficiently high. The Forward - Backward Linear Prediction (FBLP) method introduced by Nuttal [7] and Ulrich and Clayton [8] is particularly effective. The performance of the FBLP method were significantly improved by Kumaresan [9] through the Modified Forward - Backward Linear Prediction (MFBLP) method. MFBLP uses the array output correlation matrix to determine both the signal and the noise subspaces. Then the linear prediction is optimized using only the signal subspace and the resulting estimator presents improved performance at low SNR.

The proposed estimation procedure modifies the MFBLP method in three aspects. The MFBLP method optimizes the linear prediction problem through the least squares (LS) criterion. The classical least squares technique has been used to solve the overdetermined system of equations $Ax = b$, where noise is present in the data elements of A and b. When A is noise free and the noise in b is zero mean Gaussian, the LS solution is identical to maximum likelihood solution [10]. However, if A is also noisy, as in the case of the FBLP and the MFBLP methods, then the LS solution is no longer optimum. Our first modification of the MFBLP method attempts to improve the optimization criterion, substituting LS by the total least squares criterion (TLS). TLS was introduced by Golub and Van Loan [11] in 1980 to deal with this situation. Assuming that the system of equations $Ax = b$ is consistent in the absence of noise, TLS attempts to reduce the noise using perturbations ΔA and Δb in A and b such that $(A + \Delta A)x = (b+\Delta b)$ and $[\Delta A \vdots \Delta b]$ presents the smallest Frobenius norm.

The second modification results from the observation that TLS operates under the assumption that the noise components in A and b are identically independently distributed and zero mean. If there is a linear dependence among the noise components, then the TLS formulation must be reviewed to take into account the reduced dimensionality of

where $n(t) = [n_1(t) \ ... \ n_N(t)]^T$.

The output of the sensors are sampled at times t_k, $k = 1, ..., K$, leading to the noisy data vector (snapshot) y_k given by

$$y_k = Ax_k + n_k \ ; \quad k = 1, ..., K \ . \quad (8)$$

Additionally, we assume that the signal and the noise are independent zero mean complex stationary Gaussian random processes with the following second-order moments

$$E\{x_k x_l^H\} = C\delta_{k,l} \quad ; \quad E\{x_k x_l^T\} = 0$$
$$E\{n_k n_l^H\} = \sigma^2 I\delta_{k,l} \quad ; \quad E\{n_k n_l^T\} = 0 \quad (9)$$

where $E\{\}$ is the expectation, C is the unknown signal co-variance matrix, $\delta_{k,l}$ is the Kronecker delta operator, σ^2 is the unknown noise power, I is the identity matrix and $(.)^H$ denotes the conjugate transpose. These conditions imply that signal and noise at different instants are not correlated. Signal components may present some spatial correlation described by matrix C, whereas the noise components do not. Also, it can be verified that the condition $E\{x_k x_l^T\} = 0$ implies that the real and the imaginary components of the signal at the nth sensor output have equal variance. The same holds for the noise.

The concepts and equations to be presented in Sections 3 and 4 and in Appendix A refer to the kth snapshot. Then we omitted the subscript "k", like that in the symbol y_k, in order to simplify the notation in those Sections and in Appendix A.

3. FORWARD-BACKWARD LINEAR PREDICTION

Linear prediction leads to an estimator for the frequencies $\omega = [\omega_1 \ ... \ \omega_M]$ [9]. To present the estimator, consider a forward linear prediction error filter of order L ($M \le L < N$), coefficients $b^f = [-1 \ b_1^f \ ... \ b_L^f]^T$ and processing y (for the kth snapshot), as illustrated by Figure 2a, [15]. The forward prediction error is

$$e^f(n) = -y(n) + \sum_{i=1}^{L} b_i^f y(n-i); \quad n = L+1, ..., N \ . \quad (10)$$

Equation (10) can be written as $Y^f b^f = e^f$, where $e^f = [e^f(L+1) \ ... \ e^f(N)]^T$ and

$$Y^f = \begin{bmatrix} y(L+1) & \cdots & y(1) \\ \vdots & \ddots & \vdots \\ y(N) & \cdots & y(N-L) \end{bmatrix} . \quad (11)$$

Consider the polynomial $P^f(z) = -z^L + b_1^f z^{L-1} + ... + b_L^f$, obtained after the filter has been optimized by minimizing the energy of the vector e^f. In the absence of noise, $P^f(z)$ has M zeros on the unit circle (signal zeros) at the positions $\{\exp(jw_m) \ ; m = 1, ..., M\}$. The remaining $(L-M)$ zeros (noise zeros) are inside the unit circle [9], [15]. Therefore, the forward linear prediction error filter can be employed as a frequency estimator.

Consider now a backward linear prediction error filter of order L and coefficients $b^b = [-1 \ b_1^b \ ... \ b_L^b]^T$, as that of Figure 2b, processing y and producing the prediction error

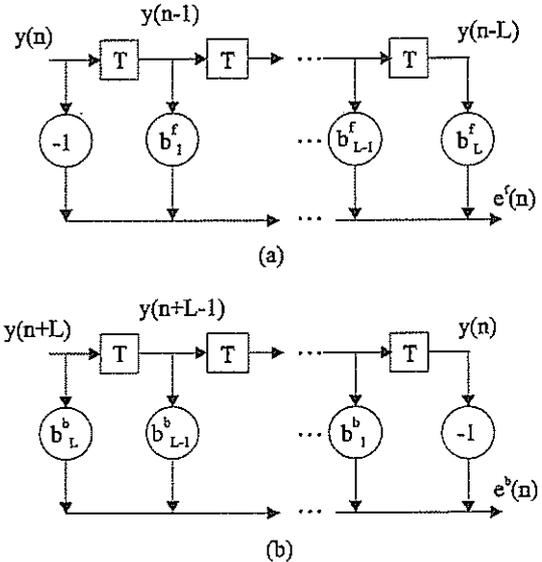$$e^b(n) = -y(n) + \sum_{i=1}^{L} b_i^b y(n+i); \quad n = 1, ..., N-L \ . \quad (12)$$

(a)

(b)

Figure 2. Linear prediction error filters: a) forward; b) backward.

Equation (12) can be written as $Y^b b^b = e^b$, where $e^b = [e^b(1) \ \cdots \ e^b(N-L)]^T$, $Y^b = Y^f J$ and J is a permutation matrix (its anti-diagonal is composed by "ones", whereas the others entries are "zeros").

After optimization, the polynomial $P^b(z) = -z^L + b_1^b z^{L-1} + ... + b_L^b$, in the absence of noise, has M zeros on the unit circle (the signal zeros) at the positions $\{\exp(jw_m) \ ; m = 1, ..., M\}$ and the remaining $(L-M)$ zeros (the noise zeros) out of the unit circle [15].

The optimum vector b_{ot}^f that minimizes the energy of e^f obeys the normal equation $(Y^f)^H Y^f b_{ot}^f = 0$, whereas the optimum vector b_{ot}^b obeys $(Y^b)^H Y^b b_{ot}^b = 0$, [15].

For stationary signals and in the limit $N \to \infty$, we can verify that $(Y^f)^H Y^f = ((Y^b)^H Y^b)^*$, implying that $b_{ot}^f = (b_{ot}^b)^*$, where $(.)^*$ denotes the complex conjugate [15].

This result motivated the definition of the forward-backward linear prediction, where $b^f = (b^b)^*$ is used even for finite N. Defining $b = b^f = (b^b)^*$, $Y = [(Y^f)^T \ (Y^b)^H]^T$ and $e = [(e^f)^T \ (e^b)^H]^T$, the formulation of the FBLP problem can now be expressed as $Yb = e$. The optimum vector b_{ot} is obtained by minimizing $\|e\|_2$, where $\|.\|_2$ denotes the 2-norm.

Consider the polynomial $P(z) = -z^L + b_1 z^{L-1} + ... + b_L$, obtained after the FBLP filter has been optimized. In the absence of noise, $P(z)$ has M zeros on the unit circle at the positions $\{\exp(jw_m) \ ; m = 1, ..., M\}$ (the signal zeros), whereas the remaining $(L-M)$ zeros (the noise zeros) are

situated inside the unit circle [9], [15]. Therefore, the frequencies ω can be estimated searching for the signal zeros on the unit circle [9].

The presence of the noise causes the zeros to fluctuate around their noise free positions. But for signal-to noise ratios (SNR) not excessively low, the frequencies ω can be estimated searching for the M zeros that are closest to the unit circle [9].

Although this property is common to the forward, backward and forward-backward arrangements, the last one presents the best performance as a frequency estimator [7], [8].

4. FBLP AND THE CONSTRAINED TO-TAL LEAST SQUARES CRITERION

The least squares criterion is usually employed to optimize the forward-backward filter. This criterion assumes that just one of the columns of Y has errors or noise, whereas, in fact, all the columns are affected by the noise. Then the total least squares (TLS) criterion [11], [16] should be employed and its application to the minimization of the energy of the FBLP error e leads to the following problem

$$\min_{(\Delta Y, b)} \|\Delta Y\|_F \quad \text{subject to} \quad (Y + \Delta Y)\, b = 0 \quad (13)$$

where $\|.\|_F$ denotes the Frobenius norm and ΔY is a matrix composed of independent variables representing perturbations to the matrix Y [11], [16].

Another aspect suggesting the use of the TLS criterion is the Toeplitz and Hankel structures of the matrix Y. Although they are not taken into account in the least squares optimization process, the Constrained Total Least Squares (CTLS) criterion is able to consider those structures [12], [16].

It is demonstrated in the appendix A that the application of the CTLS criterion to the minimization of energy of the FBLP error leads to the following problem

$$\min_{(b, \Delta y^{fb})} \left\{ \left(\Delta y^{fb} \right)^H P\, \Delta y^{fb} \right\}$$
$$\text{subject to} \quad Yb + B^H \Delta y^{fb} = 0 . \quad (14)$$

As B has rank $2(N - L)$, the solution to problem (14) is [12]

$$b_{ot} = \arg\min_b \left\{ b^H Y^H \left(B^H P^{-1} B \right)^{-1} Yb \right\} . \quad (15)$$

5. THE FBCTLS METHOD

Problem (15) refers to the kth snapshot. When all the K snapshots are considered, it is desirable to have the same b_{ot} solution for all of them, since signal and noise are stationary and DOA angles are the same for all the snapshots, implying, as shown by equations (34) and (42), in Appendix A, that both P and B matrices will also be the same for all the snapshots.

One possible strategy is to minimize the sum of the terms

74

in equation (15) for $k = 1, ..., K$, leading to

$$b_{ot} = \arg\min_b \left\{ b^H \sum_{k=1}^K \left[Y_k^H \left(B^H P^{-1} B \right)^{-1} Y_k \right] b \right\} . \quad (16)$$

It is shown in Appendix B that problem (16) can be written as

$$b_{ots} = \arg\min_b \left\{ b^H D^H Db \right\} \quad \text{subject to} \quad b_0 = -1 \quad (17)$$

where D is a matrix with dimensions $4M\,(N - L) \times (L + 1)$. This is a fourth order minimization problem with respect to b.

An iterative algorithm will now be presented to solve problem (17). In the first step we adopt $B = I$ and calculate the corresponding matrix D. Therefore, the minimization problem is reduced to a second order one which can be solved using the QR decomposition [5]. This procedure leads to a first value to b, which is used to update the B and D matrices. In the second step, a new value for b is calculated using the updated matrix D, and so on for the next steps. In general, three such steps are sufficient for an adequate convergence [5], [6].

In order to calculate the solution to the second order minimization problem at each step of the iterative algorithm, we take into account the constraint $b_0 = -1$ writing $b = \left[-1\ \eta^T \right]^T$. Then the problem at each step can be written as

$$b^H D^H Db = \|Db\|_2^2 \quad (18)$$
$$= \|-D_1 + D_2 \eta\|_2^2$$

where $D = [D_1\ D_2]$, $D_1 =$ first column of D and $D_2 \in C^{4M(N-L) \times L}$. The minimization of expression (18) leads to

$$D_2^H D_2 \eta = D_2^H D_1 . \quad (19)$$

The QR decomposition [17] produces $D_2 = QR$, with $Q \in C^{4M(N-L) \times 4M(N-L)}$, $R \in C^{4M(N-L) \times L}$ and

$$Q = \left[\ Q_1 \quad Q_2\ \right];\ Q_1 \in C^{4M(N-L) \times L} \quad (20)$$

$$R = \left[\begin{array}{c} R_0 \\ 0 \end{array} \right];\ R_0 \in C^{L \times L} . \quad (21)$$

Then we get from (19)

$$R_0^H R_0 \eta = R_0^H Q_1^H D_1 . \quad (22)$$

Since R_0 has full rank L, we finally obtain

$$\eta = R_0^{-1} Q_1^H D_1 \quad (23)$$

as the solution for each step of the iterative algorithm.

Once b_{ot} has been calculated, the polynomial $P(z)$ is formed and the corresponding zeros are calculated. The M signal zeros are estimated searching for those M zeros closest to the unit circle.

This procedure was applied to an example [6] with the following parameters: number of plane waves $M = 2$, signal correlation matrix $C = I$ (uncorrelated signals), number of sensors $N = 10$, frequencies to be estimated $\omega_1 = 0.5455$ and $\omega_2 = 0.8131$, order of the FBLP filter $L = 6$, number of

situated inside the unit circle [9], [15]. Therefore, the frequencies ω can be estimated searching for the signal zeros on the unit circle [9].

The presence of the noise causes the zeros to fluctuate around their noise free positions. But for signal-to noise ratios (SNR) not excessively low, the frequencies ω can be estimated searching for the M zeros that are closest to the unit circle [9].

Although this property is common to the forward, backward and forward-backward arrangements, the last one presents the best performance as a frequency estimator [7], [8].

4. FBLP AND THE CONSTRAINED TOTAL LEAST SQUARES CRITERION

The least squares criterion is usually employed to optimize the forward-backward filter. This criterion assumes that just one of the columns of Y has errors or noise, whereas, in fact, all the columns are affected by the noise. Then the total least squares (TLS) criterion [11], [16] should be employed and its application to the minimization of the energy of the FBLP error e leads to the following problem

$$\min_{(\Delta Y, b)} \|\Delta Y\|_F \quad \text{subject to } (Y + \Delta Y) b = 0 \quad (13)$$

where $\|.\|_F$ denotes the Frobenius norm and ΔY is a matrix composed of independent variables representing perturbations to the matrix Y [11], [16].

Another aspect suggesting the use of the TLS criterion is the Toeplitz and Hankel structures of the matrix Y. Although they are not taken into account in the least squares optimization process, the Constrained Total Least Squares (CTLS) criterion is able to consider those structures [12], [16].

It is demonstrated in the appendix A that the application of the CTLS criterion to the minimization of energy of the FBLP error leads to the following problem

$$\min_{(b, \Delta y^{fb})} \left\{ \left(\Delta y^{fb}\right)^H P \, \Delta y^{fb} \right\} \quad (14)$$
$$\text{subject to} \quad Yb + B^H \Delta y^{fb} = 0 .$$

As B has rank $2(N - L)$, the solution to problem (14) is [12]

$$b_{ot} = \arg\min_b \left\{ b^H Y^H \left(B^H P^{-1} B\right)^{-1} Yb \right\} . \quad (15)$$

5. THE FBCTLS METHOD

Problem (15) refers to the kth snapshot. When all the K snapshots are considered, it is desirable to have the same b_{ot} solution for all of them, since signal and noise are stationary and DOA angles are the same for all the snapshots, implying, as shown by equations (34) and (42), in Appendix A, that both P and B matrices will also be the same for all the snapshots.

One possible strategy is to minimize the sum of the terms

in equation (15) for $k = 1, ..., K$, leading to

$$b_{ot} = \arg\min_b \left\{ b^H \sum_{k=1}^{K} \left[Y_k^H \left(B^H P^{-1} B\right)^{-1} Y_k \right] b \right\} . \quad (16)$$

It is shown in Appendix B that problem (16) can be written as

$$b_{ots} = \arg\min_b \left\{ b^H D^H Db \right\} \text{ subject to } b_0 = -1 \quad (17)$$

where D is a matrix with dimensions $4M(N - L) \times (L + 1)$. This is a fourth order minimization problem with respect to b.

An iterative algorithm will now be presented to solve problem (17). In the first step we adopt $B = I$ and calculate the corresponding matrix D. Therefore, the minimization problem is reduced to a second order one which can be solved using the QR decomposition [5]. This procedure leads to a first value to b, which is used to update the B and D matrices. In the second step, a new value for b is calculated using the updated matrix D, and so on for the next steps. In general, three such steps are sufficient for an adequate convergence [5], [6].

In order to calculate the solution to the second order minimization problem at each step of the iterative algorithm, we take into account the constraint $b_0 = -1$ writing $b = \begin{bmatrix} -1 & \eta^T \end{bmatrix}^T$. Then the problem at each step can be written as

$$b^H D^H Db = \|Db\|_2^2 \quad (18)$$
$$= \|-D_1 + D_2 \eta\|_2^2$$

where $D = [D_1 \ D_2]$, $D_1 = $ first column of D and $D_2 \in C^{4M(N-L) \times L}$. The minimization of expression (18) leads to

$$D_2^H D_2 \eta = D_2^H D_1 . \quad (19)$$

The QR decomposition [17] produces $D_2 = QR$, with $Q \in C^{4M(N-L) \times 4M(N-L)}$, $R \in C^{4M(N-L) \times L}$ and

$$Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} ; \ Q_1 \in C^{4M(N-L) \times L} \quad (20)$$

$$R = \begin{bmatrix} R_0 \\ 0 \end{bmatrix} ; \ R_0 \in C^{L \times L} . \quad (21)$$

Then we get from (19)

$$R_0^H R_0 \eta = R_0^H Q_1^H D_1 . \quad (22)$$

Since R_0 has full rank L, we finally obtain

$$\eta = R_0^{-1} Q_1^H D_1 \quad (23)$$

as the solution for each step of the iterative algorithm.

Once b_{ot} has been calculated, the polynomial $P(z)$ is formed and the corresponding zeros are calculated. The M signal zeros are estimated searching for those M zeros closest to the unit circle.

This procedure was applied to an example [6] with the following parameters: number of plane waves $M = 2$, signal correlation matrix $C = I$ (uncorrelated signals), number of sensors $N = 10$, frequencies to be estimated $\omega_1 = 0.5455$ and $\omega_2 = 0.8131$, order of the FBLP filter $L = 6$, number of

snapshots $K = 100$. One hundred different experiments were produced with the above specifications, but with different realizations for the random signal and the noise. Referring to expression (9), the signal-to-noise ratio (SNR) is defined as SNR= $1/\sigma^2$.

Figure 3a) presents the superposition of the zeros of the polynomial $P(z)$ for the 100 experiments and SNR= 15 dB. There are two signal zeros close to the ideal positions defined by the desired frequencies on the unit circle, whereas the noise zeros are distributed inside the unit circle. Therefore, the desired frequencies can be estimated searching for the two zeros that are closest to the unit circle.
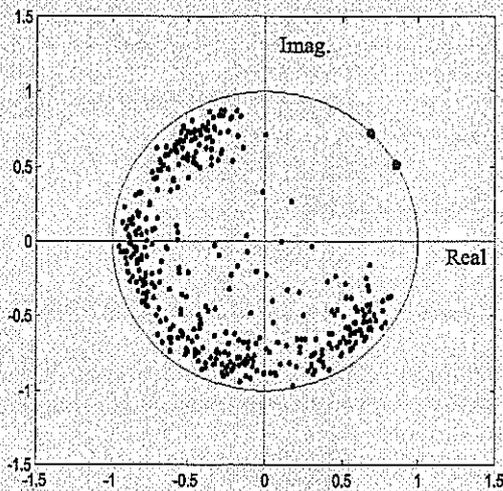
(a)

(b)

Figure 3. Zeros in the z-plane: a) SNR=15 dB; b) SNR=0 dB.

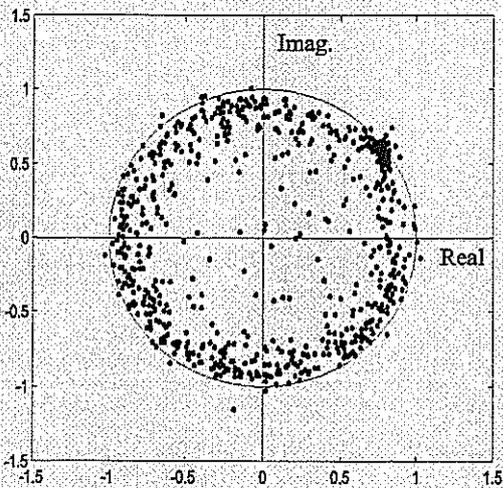However, the variance of the zeros increases as the SNR decreases and for SNR $<$ 5 dB there are noise zeros closer to unit circle than the signal zeros, as shown in Figure 3b. Therefore, the estimation strategy based on searching for the two zeros that are closest to the unit circle leads to poor performance for small values of the SNR.

We adopted another criterion for the selection of the signal zeros to overcome this problem. We select the M signal zeros as those that minimize the maximum likelihood cost function for the DOA problem, given by [6]

$$J(\omega) = \left[A \left(A^H A \right)^{-1} A^H - I \right] \left(\sum_{k=1}^{K} y_k y_k^H (k) \right) \quad (24)$$

where A is formed by the arguments of the zeros to be tested. For each group of M zeros among the L zeros obtained at each experiment, we form the corresponding A matrix and calculate $J(\omega)$.

Figure 4 shows the results achieved with the proposed procedure applied to the example with $L = 7$. This value of L maximizes the performance. The root mean square error is calculated for both frequencies after 100 different experiments and for various values of the SNR. Figure 4 also presents the performance of the methods MODE [5] and MODEX [6].

FBCTLS presents better performance than the classical MODE for SNR values smaller than 5 dB and is competitive with MODEX.
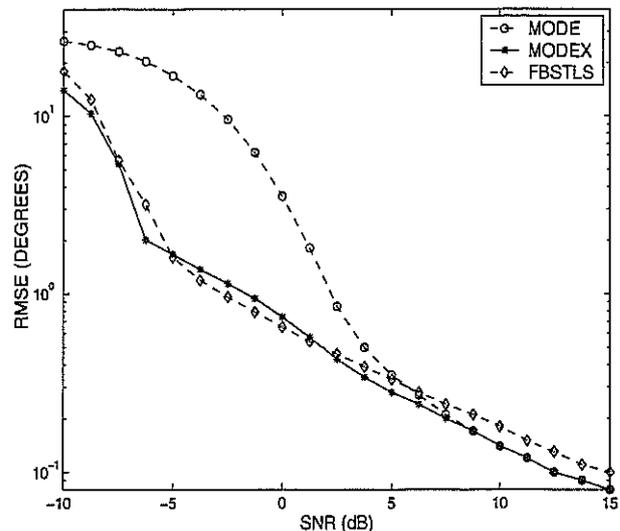
Figure 4. Root mean square error as a function of the SNR for FBCTLS, MODE and MODEX.

In order to compare the computational complexity of the three methods, we observe that MODE and MODEX solve problems similar to (17). But MODEX and FBCTLS use $L > M$, whereas MODE uses $L = M$. So MODE is the most economical in this part of the algorithm. MODE is also the most economical when searching for the signal zeros because it uses $L = M$, that is, just the signal zeros. On the other hand, MODEX and FBCTLS use the maximum likelihood cost function in (24) to select the signal zeros. Therefore, MODE is the most economical in terms of computational complexity.

The comparison between MODEX and FBCTLS is not an easy task. MODEX search for the signal zeros in a ensemble composed by $(L + M)$ zeros, while FBCTLS uses L zeros. Thus MODEX is less economical than FBCTLS as it demands additional efforts to execute the MODE and to search

for the signal zeros. On the other hand, the matrices in problem (17) for the FBCTLS have larger dimensions then those related to MODEX. Therefore, a detailed quantitative analysis is necessary to compare the numerical complexity of both methods, but that is beyond the scope of this paper.

6. CONCLUSIONS

We proposed a DOA method, FBCTLS, based on forward-backward linear prediction and the Constrained Total Least Squares criterion.

The proposed procedure leads to a minimization problem similar to that presented by some popular maximum likelihood methods. The solution was achieved by means of an iterative algorithm similar to the one employed in the MODE and MODEX methods.

Numerical comparisons showed that FBCTLS and MODEX perform similarly and that both are better than MODE. However, both demand additional computational effort when compared to MODE.

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, July 1996.

[2] S. M. Kay, *Modern Spectral Estimation - Theory and Application*, Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1988.

[3] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-34, no. 5, pp. 1081–1089, Oct. 1986.

[4] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *Proc. Inst. Elect. Eng.*, vol. 137, no. 1, pp. 19–26, Feb. 1990.

[5] J. Li, P. Stoica, and Z. Liu, "Comparative study of IQML and MODE direction-of-arrival estimators," *IEEE Trans. Signal Processing*, vol. 46, no. 1, pp. 149–160, Jan. 1998.

[6] A. B. Gershman and P. Stoica, "New MODE-based techniques for direction finding with an improved threshold performance," *Signal Processing*, , no. 76, pp. 221–235, 1999.

[7] A. H. Nuttal, "Spectral analysis of a univariate process with bad data points via maximum entropy and linear predictive techniques," *Naval Underwater System Center (NUSC) scientific and engineering studies, Spectral Estimation, NSUC, New London, CT*, Mar. 1976.

[8] T. J. Ulrych and R. W. Clayton, "Time series modelling and maximum entropy," *Phys. Earth Planet. Interiors*, vol. 12, pp. 188–200, Aug. 1976.

[9] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood," *Proc. of the IEEE*, vol. 70, pp. 975–989, Sept. 1982.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1993.

[11] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM Journal on Numerical Analysis*, vol. 17, pp. 883–893, 1980.

[12] T. J Abatzoglou, J. M. Mendel, and G. A. Harada, "The constrained total least squares technique and its application to har-

monic superresolution," *IEEE Trans. Signal Processing*, vol. 39, no. 5, pp. 1070–1087, May 1991.

[13] S. Haykin, editor, *Advances in Spectrum Analysis and Array Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1991.

[14] S. Unnikrshina Pillai, *Array Signal Processing*, Springer-Verlag, New York, 1989.

[15] S. Haykin, *Adaptive Filter Theory*, Prentice Hall Information and System Sciences Series, Englewood Cliffs, NJ, 2nd edition, 1991.

[16] R. P. Lemos, *Mnimos-quadrados totais e mxima verossimilhana em estimao de freqncias*, Ph.D. thesis, School of Electrical and Computer Engineering - UNICAMP, Campinas-SP-Brazil, June 1997.

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins Univ. Press, Baltimore, 2nd edition, 1996.

[18] R. P. Lemos and A. Lopes, "A unifying framework to total least squares and approximate maximum likelihood," in *Bi-Annual Intern. Telecom. Symposium - ITS'96*, Acapulco-Mexico, Oct. 1996, pp. 153–157.

7. APPENDICES

A. SOLUTION TO THE CTLS PROBLEM

The application of the CTLS criterion to the minimization of the energy of the FBLP error leads to following problem [12]:

$$\min_{(\Delta Y, b)} \|\Delta Y\|_F \qquad (25)$$

subject to $\quad (Y + \Delta Y) b = 0 \qquad (26)$

where

$$\Delta Y = \begin{bmatrix} \Delta Y^f \\ (\Delta Y^b)^* \end{bmatrix} \qquad (27)$$

$$\Delta Y^f = \begin{bmatrix} F_1^f \Delta y & \dots & F_{L+1}^f \Delta y \end{bmatrix} \qquad (28)$$

$$\Delta Y^b = \begin{bmatrix} F_1^b \Delta y & \dots & F_{L+1}^b \Delta y \end{bmatrix} \qquad (29)$$

$$\Delta y = [\Delta y(1) \ \dots \ \Delta y(N)]^T . \qquad (30)$$

The vector Δy is composed of independent variables representing perturbations to the vector y. The matrices

$$F_i^f \in \mathcal{R}^{(N-L) \times N} \quad \text{and} \quad F_i^b \in \mathcal{R}^{(N-L) \times N}$$
$$i = 1, \dots, L+1 \qquad (31)$$

are composed by "ones" and "zeros" describing the structures of the matrices Y^f and Y^b, respectively. Each row has $(N - 1)$ "zeros" and just one "one", whose position is chosen to assure that

$$Y^f = \begin{bmatrix} F_1^f y & \dots & F_{L+1}^f y \end{bmatrix} \qquad (32)$$

and

$$Y^b = \begin{bmatrix} F_1^b y & \dots & F_{L+1}^b y \end{bmatrix} . \qquad (33)$$

The matrices ΔY^f and ΔY^b are composed by the perturbation variables of Δy and present the same structure of Y^f and Y^b, respectively.

Referring to $\|\Delta Y\|_F$ in problem (25) and using expressions (27), (28), (29) and (30), it is possible to write

$$\|\Delta Y\|_F^2 = \left(\Delta y^{fb}\right)^H \mathbf{P} \; \Delta y^{fb} \tag{34}$$

where

$$\Delta y^{fb} = \left[\begin{array}{c} \Delta y \\ \Delta y^* \end{array} \right]_{2N \times 1} \tag{35}$$

$$\mathbf{P} = diag\left[p_1 \; ... \; p_N \; p_1 \; ... \; p_N\right]_{2N \times 2N} \tag{36}$$

and the real numbers p_i, $(i = 1, ..., \; N)$ are weighting factors that take into account how many times $\Delta y(i)$ and $\Delta y(i)^*$ appear in ΔY.

Referring to equation (26), it is possible to verify that

$$\Delta Y^f b = \left[-F_1^f + \sum_{i=2}^{L+1} F_i^f b_{i-1}\right] \Delta y \tag{37}$$

$$\left(\Delta Y^b\right)^* b = \left[-F_1^b + \sum_{i=2}^{L+1} F_i^b b_{i-1}\right] \Delta y^* .$$

Let

$$\mathbf{B}_f = \left[-F_1^f + \sum_{i=2}^{L+1} F_i^f b_{i-1}\right]^H_{(N-L) \times N}$$

$$\mathbf{B}_b = \left[-F_1^b + \sum_{i=2}^{L+1} F_i^b b_{i-1}\right]^H_{(N-L) \times N} . \tag{38}$$

Then

$$\mathbf{B}_f = \left[\begin{array}{cccccccc} b_L & b_{L-1} & \cdots & b_1 & -1 & 0 & \ldots & 0 \\ 0 & b_L & \ldots & b_2 & b_1 & -1 & \ldots & 0 \\ & \ddots & & \ddots & & \ddots & & \\ 0 & 0 & \ldots & & \ldots & & \ldots & -1 \end{array} \right]^H \tag{39}$$

$$\mathbf{B}_b = \mathbf{J}\mathbf{B}_f\mathbf{J} \tag{40}$$

where $\mathbf{J}$ is a permutation matrix.

Using these matrices and expression (27), it is possible to write

$$(\Delta Y) b = \left[\begin{array}{cc} \mathbf{B}_f^H & 0 \\ 0 & \mathbf{B}_b^H \end{array} \right] \left[\begin{array}{c} \Delta y \\ \Delta y^* \end{array} \right] . \tag{41}$$

Let

$$\mathbf{B}^H = \left[\begin{array}{cc} \mathbf{B}_f^H & 0 \\ 0 & \mathbf{B}_b^H \end{array} \right]_{2(N-L) \times 2N} . \tag{42}$$

Using (41), (42) and (35), expression (26), can be written as

$$\mathbf{Y}b + \mathbf{B}^H \Delta y^{fb} = 0 . \tag{43}$$

Finally, using expressions (34) and (43) in problem (25) results

$$\min_{(b, \Delta y^{fb})} \left\{ \left(\Delta y^{fb}\right)^H \mathbf{P} \; \Delta y^{fb} \right\} \tag{44}$$

$$\text{subject to} \quad \mathbf{Y}b + \mathbf{B}^H \Delta y^{fb} = 0 . \tag{45}$$

B. TRANSFORMATION OF THE CTLS PROBLEM

Problem (16) is repeated here for convenience:

$$b_{ot} = \arg\min_{b} \left\{ b^H \sum_{k=1}^{K} \left[Y_k^H \left(\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B}\right)^{-1} Y_k \right] b \right\} . \tag{46}$$

We will introduce some modifications in equation (46). It can be verified that

$$Y_k b = \mathbf{B}^H y_k^{fb} . \tag{47}$$

Applying expression (47) into expression (46) yields

$$b_{ot} =$$
$$= \arg\min_{b} \left\{ \sum_{k=1}^{K} \left[\left(y_k^{fb}\right)^H \mathbf{B} \left(\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B}\right)^{-1} \mathbf{B}^H y_k^{fb} \right] \right\}$$
$$= \arg\min_{b} \; trace \left\{ \mathbf{B} \left(\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B}^H\right)^{-1} \mathbf{B}^H \widehat{R}_y^{fb} \right\} \tag{48}$$

where

$$\widehat{R}_y^{fb} = \sum_{k=1}^{K} y_k^{fb} \left(y_k^{fb}\right)^H \tag{49}$$

is a correlation estimator, except for a normalizing constant.

Problem (48) is the same as those obtained in [5], [6], [3] and [18], where the maximum likelihood criterion is used for the estimation of the frequencies ω.

Now a subspace restriction is used to reduce the effect of the noise in the $\widehat{R}_y^{fb}$ matrix, following a procedure inspired by the one presented in [5].

Consider the singular value decomposition of the matrix $\widehat{R}_y^{fb}$

$$\widehat{R}_y^{fb} = \mathbf{U}\Sigma\mathbf{U}^H \tag{50}$$

where

$$\mathbf{U} = [u_1 \; ... \; u_{2N}] \; ; \quad \mathbf{U}^H\mathbf{U} = \mathbf{I}$$

$$\Sigma = diag\left(\lambda_1 \; ... \; \lambda_{2N}\right) \tag{51}$$

and u_i are the singular vectors and λ_i are the singular values ordered from the largest one to the smallest one [15].

In the case $n_k = 0$ and for $K \to \infty$, it is shown in [15] that $\lambda_i = 0$ for $2M + 1 \leq i \leq 2N$. Imposing this condition on $\widehat{R}_y^{fb}$ of (49) produces

$$\widehat{R}_y^{fbM} = \mathbf{U}_S\Sigma_S\mathbf{U}_S^H \tag{52}$$

where

$$\mathbf{U}_S = [u_1 \; ... \; u_{2M}] \; ; \quad \Sigma_S = diag\left(\alpha_1 \; ... \; \alpha_{2M}\right)$$

$$\alpha_i = \frac{1}{\lambda_i} \left(\lambda_i - \frac{\sum_{n=2M+1}^{2N} \lambda_n}{2N - 2M} \right)^2 \; ; \; i = 1, ..., 2M . \tag{53}$$

This modification aims to reduce the effect of the noise on $\widehat{R}_y^{fb}$, by eliminating the participation of the noise subspace.

The values adopted for α_i, $1 \leq i \leq 2M$, reduce the effect of the noise on the remaining singular values λ_i belonging to the signal subspace [6].

Using the results of (52) into expression (48) results

$$\mathbf{b}_{ot} =$$
$$= \arg\min_{\mathbf{b}} \; trace \left\{ \mathbf{B} \left(\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{U}_S \mathbf{\Sigma}_S \mathbf{U}_S^H \right\} \quad (54)$$

Let $\mathbf{U}_S \left(\mathbf{\Sigma}_S \right)^{0.5} = \mathbf{V}$. It can be verified that $\mathbf{B}^H \mathbf{V} = [\mathbf{S}_1 \mathbf{b} \; ... \; \mathbf{S}_{2M} \mathbf{b}]$ where

$$\mathbf{S}_i = \begin{bmatrix} v_{L+1,i} & \cdots & v_{1,i} \\ \vdots & \ddots & \vdots \\ v_{N,i} & \cdots & v_{N-L,i} \\ v_{N+1,i} & \cdots & v_{N+L+1,i} \\ \vdots & \ddots & \vdots \\ v_{2N-L,i} & \cdots & v_{2N,i} \end{bmatrix} . \quad (55)$$

Substituting these results into equation (54) yields

$$\mathbf{b}_{ot} =$$
$$= \arg\min_{\mathbf{b}} \; trace \left\{ \left(\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{V} \mathbf{V}^H \mathbf{B} \right\}$$
$$= \arg\min_{\mathbf{b}} \; trace \left\{ [\mathbf{S}_1 \mathbf{b} \; ... \; \mathbf{S}_{2M} \mathbf{b}]^H \left(\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B} \right)^{-1} \right.$$
$$\left. [\mathbf{S}_1 \mathbf{b} \; ... \; \mathbf{S}_{2M} \mathbf{b}] \right\} . \quad (56)$$

Considering the Cholesky decomposition

$$\mathbf{B}^H \mathbf{P}^{-1} \mathbf{B} = \mathbf{G}^H \mathbf{G} \quad (57)$$

problem (56) can be written as

$$\mathbf{b}_{ot} = \arg\min_{\mathbf{b}} \; trace \left\{ [\mathbf{S}_1 \mathbf{b} \; ... \; \mathbf{S}_{2M} \mathbf{b}]^H \right.$$
$$\left. \mathbf{G}^{-1} \left(\mathbf{G}^H \right)^{-1} [\mathbf{S}_1 \mathbf{b} \; ... \; \mathbf{S}_{2M} \mathbf{b}] \right\} . \quad (58)$$

But

$$\left(\mathbf{G}^H \right)^{-1} [\mathbf{S}_1 \mathbf{b} \; ... \; \mathbf{S}_{2M} \mathbf{b}] =$$
$$= \left[\left(\mathbf{G}^H \right)^{-1} \mathbf{S}_1 \mathbf{b} \; ... \; \left(\mathbf{G}^H \right)^{-1} \mathbf{S}_{2M} \mathbf{b} \right] \quad (59)$$

and

$$trace \left\{ \left[\left(\mathbf{G}^H \right)^{-1} \mathbf{S}_1 \mathbf{b} \; ... \; \left(\mathbf{G}^H \right)^{-1} \mathbf{S}_{2M} \mathbf{b} \right]^H \right.$$
$$\left. \left[\left(\mathbf{G}^H \right)^{-1} \mathbf{S}_1 \mathbf{b} \; ... \; \left(\mathbf{G}^H \right)^{-1} \mathbf{S}_{2M} \mathbf{b} \right] \right\} =$$
$$= \sum_{m=1}^{2M} \left[\left(\mathbf{G}^H \right)^{-1} \mathbf{S}_i \mathbf{b} \right]^H \left[\left(\mathbf{G}^H \right)^{-1} \mathbf{S}_i \mathbf{b} \right] = \quad (60)$$
$$= \mathbf{b}^H \mathbf{D}^H \mathbf{D} \mathbf{b}$$

where

$$\mathbf{D} = \begin{bmatrix} \left(\mathbf{G}^H \right)^{-1} \mathbf{S}_1 \\ \vdots \\ \left(\mathbf{G}^H \right)^{-1} \mathbf{S}_{2M} \end{bmatrix}_{4M(N-L)\times(L+1)} . \quad (61)$$

Using (61) into problem (58), we have

$$\mathbf{b}_{ot} = \arg\min_{\mathbf{b}} \left\{ \mathbf{b}^H \mathbf{D}^H \mathbf{D} \mathbf{b} \right\} \quad \text{subject to} \quad b_0 = -1 . \quad (62)$$

The minimization in problem (62) is subjected to $b_0 = -1$ in agreement with the definition of vector b.

Amauri Lopes received the B.S. , the M. Sc. and the Ph.D. degrees in Electrical Engineering from the University of Campinas-SP-Brazil, in 1972, 1974 and 1982, respectively. Since 1973 he has been with the Electrical and Computer Engineering School, University of Campinas, where he is currently an associate professor. His research areas are digital signal processing, circuit theory and digital communications.

Ivanil Sebastião Bonatti was born in Mogi-Mirim - SP, Brazil, in 1951. He received the B.Sc. (1973) and M.Sc. (1975) degrees in Electrical Engineering from the University of Campinas, UNICAMP, and the "Doctorat en Automatique" degree from the University Paul Sabatier, Toulouse, France, in 1981. He joined the School of Electrical and Computer Engineering of the University of Campinas in 1975, where he is currently a professor. His main interests are circuit theory, simulation analysis and telecommunications.

Pedro Luis Dias Peres was born in Sorocaba - SP, Brazil, in 1960. He received the B.Sc. and M.Sc. degrees in Electrical Engineering from the University of Campinas, UNICAMP, in 1982 and 1985 respectively, and the "Doctorat en Automatique" degree from the University Paul Sabatier, Toulouse, France, in 1989. In 1990 he joined the School of Electrical and Computer Engineering of the University of Campinas, Brazil, where he is currently an associate professor. His main interests are control theory, convex analysis and optimization and circuit theory.

Ricardo Fialho Colares was born in Fortaleza, Ceará, Brazil, in 1969. He received the B.S. Degree in Electrical Engineering from the Universidade Federal do Ceará, Brazil, in 1992. The M.Sc. and the Ph.D. Degrees in Electrical Engineering was received from the University of Campinas (UNICAMP), SP, Brazil, in 1997 and 2000, respectively. He has participated in a research program with the CPqD-UNICAMP along 1998 and 1999. During 2000 he has joint to the Universidade São Francisco, SP, Brazil, as Undergraduate Professor. Since the beginning of 2001 he has been with Universidade de Fortaleza (UNIFOR), CE, Brazil, where he is currently a Titular Professor and Coordinator of the Undergraduate Telecommunication Engineering Course. His research interests areas are digital signal processing, applied mathematics in telecommunications and antenna array.

Carlos Antonio Alves received the B.S. and M. Sc. degrees in Electrical Engineering from the São Paulo State University - UNESP, Ilha Solteira-SP- Brazil, in 1989 and 1995, respectively. Since 1990 he has been with the Department of Electrical Engineering of the School of Engineering of Ilha Solteira - FEIS, at UNESP where is an Assistant Professor. Currently he has been pursuing the Ph. D. degree at the School of Electrical and Computer Engineering in the University of Campinas - UNICAMP. His research area is digital signal processing.