ON MODEL ORDER DETERMINATION OF COMPLEX EXPONENTIAL SIGNALS

Chinghui J. Ying and Randolph L. Moses

Department of Electrical Engineering, The Ohio State University, 2015 Neil Avenue, Columbus, OH 43210

Abstract. We present an algorithm for model order determination and corresponding maximum likelihood parameter estimation for complex exponential signal modeling. The algorithm exploits initial nonparametric frequency location estimates to significantly reduce the search space for the correct model order and parameter estimates. An FFT is used to obtain initial frequency region estimates. An initial overdetermined model order and initial frequency estimates are obtained using the CRB resolution limit and the FFT peaks. After iterative minimization, a post-processing step eliminates the extraneous sinusoidal modes using CRB resolution limits and statistical detection tests. Because the algorithm searches on only a limited set of model orders and parameter regions, it is computationally tractable even for large data lengths and model orders. Simulations are provided to illustrate the performance of the proposed algorithm.

Key Words. Model Order Determination, Exponential Modeling, Detection

1. INTRODUCTION

Exponential modeling arises in many areas, including speech processing, deconvolution, radar and sonar signal processing, array processing, and spectrum estimation (Parthasarathy and Tufts, 1987; Steedly and Moses, 1991; Lee *et al.*, 1990; Cadzow, 1982). There are two main research issues involved in the exponential modeling problem. One is the determination of the model order, *i.e.*, the number of modes in the signal. The second question is the estimation of the model parameters.

Model order determination using information-theoretic criteria has been addressed in (Yin and Krishnaiah, 1987; Wax and Kailath, 1985; Wax and Ziskind, 1989; Wu and Fuhrmann, 1991). These algorithms search the entire model order space in order to find the correct model order according to some information criterion (AIC, MDL, etc.). For nonparametric algorithms (*e.g.*, Wax and Kailath (1985)), the computation cost is modest but performance suffers even for moderate signal to noise ratios (SNRs). In contrast, parametric algorithms (Wax and Ziskind, 1989; Wu and Fuhrmann, 1991) give good results but are computationally expensive.

Several authors have used singular value analysis in the detection problem (Cadzow, 1982; Rao and Gnanaprakasam, 1988). The basic idea there is to identify the dominant singular values (SVs) by some criterion (for instance, a pre-chosen threshold). The number of the dominant SVs is the model order. Rao and Gnanaprakasam (1988) developed a statistical criterion to determine the threshold. Bakamidis *et al.* (1991) developed a technique based on the synthesis of the noiseless signal using the principal component analysis and the SVs. The fitted error power is then compared to the assuming known noise power. The best fit (in terms of the estimate of noise power) gives the best noiseless signal estimate. The number of SVs used in the synthesis is then chosen to be the estimate of the correct model order.

Most of these methods become computationally unwieldy if the data length and model order become large. For example, the SVD-based methods involve computing the SVD of a matrix whose minimum dimension is approximately $N/3$, where N is the data length. For large N ($N = 1024$, for example) this decomposition is computationally expensive. The parametric order estimation methods (Wax and Ziskind, 1989; Wu and Fuhrmann, 1991) essentially require testing all model orders up to a maximum order which is related to the data length. This requires a large number of (iterative) nonlinear minimizations, and is also computationally prohibitive. For such applications, the computational overhead makes these algorithms difficult or impossible to use in practice.

In this paper we propose a computationally tractable alternative for model order determination. The algorithm uses an initial nonparametric frequency estimate, obtained using an FFT, to reduce considerably the set of candidate model order hypotheses. In addition, we significantly reduce the initial condition points for nonlinear minimization procedures by use of an FFT initialization which is combined with a mode splitting algorithm to account for the limited FFT resolution. The splitting algorithm is based on the CRB resolution limit. Because the initialization procedure tends to initially overestimate the model order, we include a post-processing step to test for and eliminate extraneous modes in the model. This process of intelligently searching for model order and parameter estimates in a reduced subset of the entire space results in significant computational savings.

An outline of this paper is as follows. In Section 2 we present the data model and the problem formulation. In Section 3 we discuss the statistical background for the proposed model order determination and derive the proposed algorithm. In Section 4 we present simulation studies which demonstrate the ability of the proposed algorithm. Finally, in Section 5 we conclude the paper.

2. DATA MODEL AND PROBLEM FORMULATION

Consider a data vector $\{y_q\}_{q=0}^{m-1}$ of complex exponential samples

$$y_q = \sum_{i=1}^{n} \alpha_i e^{j(\omega_i q + \phi_i)} + e_q \tag{1}$$

where $\alpha_i \in R$, ω_i, $\phi_i \in (-\pi,\ \pi]$. We assume that the exponential modes are all distinct, *i.e.*, $\omega_k \neq \omega_l$ for $k \neq l$. We also assume that $\{e_q\}$ is a zero mean complex white Gaussian noise sequence with variance σ^2.

There are two main questions in exponential modeling. One is the estimation problem, *i.e.*, to estimate the parameters, $\{\alpha_i\}_{i=1}^n$, $\{\phi_i\}_{i=1}^n$, and $\{\omega_i\}_{i=1}^n$ in the model. The other problem is to determine the number of modes in the signal. In this paper we consider a combined parameter/order estimation algorithm.

To address the order determination problem, we use results of maximum likelihood (ML) estimation of the parameter vector:

$$\theta = \left[\ \{\alpha_i\}_{i=1}^n \quad \{\phi_i\}_{i=1}^n \quad \{\omega_i\}_{i=1}^n\ \right]^T . \tag{2}$$

Under the assumptions made, the negative log-likelihood function for an estimate $\widehat{\theta}$ of θ is given as (Bresler and Macovski, 1986; Stoica *et al.*, 1989)

$$L\left(\widehat{\theta}\right) = \sum_{q=0}^{m-1} \left| y_q - x_q\left(\widehat{\theta}\right)\right|^2 , \tag{3}$$

where $x_q\left(\widehat{\theta}\right)$ is an estimate of the noiseless version of y_q parameterized by $\widehat{\theta}$. The ML estimate is then given by

$$\widehat{\theta}_{ML} = \arg\min_{\widehat{\theta}} L\left(\widehat{\theta}\right) . \tag{4}$$

Note that the ML estimate given in Equation 4 is only valid if the number of modes in the estimate is correct, *i.e.*, $\hat{n} = n$. By increasing the number of modes in the estimate, we can always reduce the loss function $L\left(\widehat{\theta}\right)$.

We design an algorithm which obtains the correct model order and the corresponding ML parameter estimates. We first obtain initial estimates using FFT peaks. We then apply a splitting algorithm that is based on a criterion derived from the CRB results to refine the initial estimates. We next apply an ML procedure (using Equation 4) to these refined initial estimates. In the computer simulations provided, the ML estimation using artificial neural networks in Ying *et al.* (1993) is used; however, other methods (*e.g.*, Stoica *et al.* (1989) or Ziskind and Wax (1988)) could be used. After the ML step, we use a pruning algorithm to prune the extra modes in the estimate. We repeat the ML estimation and the pruning cycle until no mode is pruned. The number of modes in the final estimate is the estimate of the model order; the final estimate is also the ML estimate of the model parameters.

We note that the above proposed algorithm estimates the model order and model parameters simultaneously. Moreover, the algorithm does not search the entire space for the possible model order. The algorithm only considers a "small range" of the entire space. Also, it is shown by the computer simulations that the algorithm provides a consistent estimate of the model order (as SNR $\to \infty$ or data length $m \to \infty$).

3. DERIVATION OF THE PROPOSED ALGORITHM

As described in the previous section, the proposed model order determination algorithm is based on the FFT peaks of the given signal. However, the FFT has resolution limits and thus cannot detect (and estimate) closely spaced modes. To solve this resolution limitation problem, we design an algorithm to split the FFT peaks into some larger

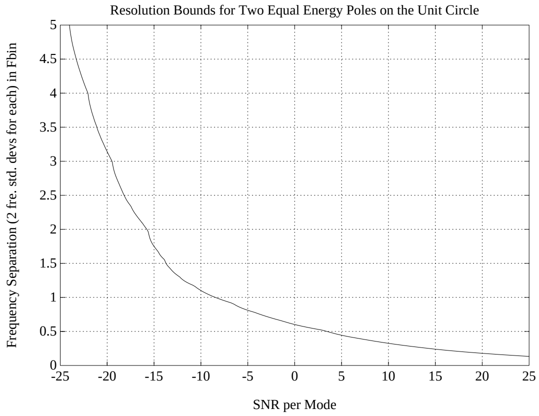

Fig. 1. Resolution Limits for data length $m = 25$.

number of peaks based on certain resolution limit obtained from the CRB. This splitting algorithm introduces additional modes. Consequently, we increase the model order in this step, and sometimes overestimate the model order. To prune any extra modes, we design a pruning algorithm which is also based on the resolution limit.

3.1. *RESOLUTION LIMIT*

The resolution limit concerns the question of how close (in frequency) the modes can be for an algorithm to resolve them. For any unbiased estimator, the Cramér-Rao Bound (CRB) gives the best performance one can obtain. We thus define the resolution limit u for any two poles on the unit circle, $p_1 = e^{j\omega_1}$ and $p_2 = e^{j\omega_2}$, as

$$u \triangleq c\left(\sigma_{\omega_1} + \sigma_{\omega_2}\right), \tag{5}$$

where the angle CRBs of p_1 and p_2 are $\sigma^2_{\omega_1}$ and $\sigma^2_{\omega_2}$, respectively (the CRB expressions can be found, for example, in Lee (1992) or Steedly and Moses (1993)). The parameter c is selected by the user, and controls the confidence with which two modes can be resolved; we use $c = 2$ which corresponds to 95% confidence intervals for the frequency estimates of the modes. That is, when two poles are at this limit, the 95% confidence intervals of the frequency estimates for each pole become disjoint. Note that this limit gives a lower bound for all unbiased estimators since it is based on the CRB.

To give an example of how the resolution limits work, we plot the resolution bounds for data length $m = 25$ in Figure 1[1]. From the figure we can see that given SNR/mode=0 dB, the resolution limit is about 0.65 Fourier separation bins (denoted Fbin) at data length $m = 25$. In another words, if any two modes are less than 0.65 Fbin apart at SNR/mode=0 dB and $m = 25$, no unbiased estimators can resolve them statistically (namely, more than 5% of the two mode estimates will be overlapped).

Note that two modes of equal energy give the best resolvability among all pairs of modes; for unequal energy cases, the resolution limits are larger than the one for the equal energy case. This can be seen from the CRB expression. In fact, from Equation 5, if one of the CRBs becomes larger, the resolution limit u consequently becomes larger. In addition, the two-mode resolution bound is a lower bound for frequency resolution of multiple modes. If three or more modes are closely spaced, the CRB resolution limit is higher than the two mode limit (Lee, 1992); thus, the limit shown in Figure 1 gives a lower bound on the resolution limit.

The maximum number of modes within a Fbin is inversely proportional to the resolution limit. In fact, the maximum number of modes in one Fbin is defined as

$$v = \lceil \frac{1}{u} \rceil, \tag{6}$$

where $\lceil a \rceil$ denotes the smallest integer which is larger than a.

To compute the resolution limit u, we only need the data length m and the noise power σ^2. For the unknown noise power case, one can use the approach in Stoica *et al.* (1992) to obtain an estimate of the noise power $\widehat{\sigma^2}$.

3.2. *SPLITTING ALGORITHM*

The purpose of this splitting algorithm is to overcome the resolution limitation problem inherent from FFT. It is well known that FFT cannot resolve modes which are less than 1 Fbin apart. If a windowed FFT is used, the resolution ability degrades by an amount that depends on the window (Harris, 1978). In order to obtain "high resolution" results, we introduce additional modes. Consequently, we increase the model order in this step, and sometimes overestimate the model order.

As described above, for any given data length m and noise power σ^2, we can calculate the maximum number of modes in one Fbin, v. From the peaks of the FFT, we obtain the initial estimates for the parameters. We then use the maximum number of modes in one Fbin to split each of the FFT peaks into v peaks. We keep the same magnitudes and phases for split peaks but change

[1] Here, two equal energy modes are used. The bounds shown are the lower bounds for two equal energy modes and data length $m = 25$ over all initial phases. That is, $\sigma_{\omega_i} = \arg\min_\phi \sigma_{\omega_i}(\phi)$.

the frequencies. For instance, for $m = 25$ and SNR/mode=0 dB, the resolution limit u is approximately 0.65 Fbin. Thus, the maximum number of modes per Fbin is $v = \lceil \frac{1}{0.65} \rceil = \lceil 1.538 \rceil = 2$. Given this information, we split a peak, say at the 5th Fbin, into two peaks located at frequencies $5 \pm \frac{1}{2}u = 5.325, 4.675$. These become our frequency estimates after splitting.

The splitting algorithm gives the maximum number of modes in the signal that can be resolved, *i.e.*, the possible maximum model order. This ensures the model order is not underestimated (with a statistical confidence that depends on c in Equation 5), and is likely overestimated. Our subsequent processing is aimed at eliminating any extraneously modeled modes. This is done by the pruning algorithm described below.

3.3. *PRUNING ALGORITHM*

In the splitting algorithm we often introduce extraneous modes in the estimate. In this step we prune these extraneous modes by a resolution test and model complexity test.

After the ML estimation step, any extraneous modes will fall into the following two situations:

1. Two or more estimated modes can correspond to a single true mode. In this case the modes are, in general, within each other's resolution limit and can be combined. The combining job is done as follows.
   1a. Calculate the CRBs and the resolution limits for the estimates from the ML estimation. Note that to be able to calculate the CRBs, we need to estimate the noise power σ^2 (*e.g.*, Stoica *et al.* (1992)).
   1b. Discard a lower energy mode if it is within the resolution limit of some larger energy mode. Combine only one mode at a time.
2. The magnitudes of the extra modes will be small. In this case the extra modes are used to model the noise, and we eliminate these extra modes by setting a threshold on the power for a given SNR. For example, it is reasonable to set the lower bound to be the average noise level. Modes whose powers are smaller than this threshold are eliminated. We have the following empirical calculation for the threshold,

$$\rho_l = 10^{-(\sqrt{m}+\widehat{\mathrm{SNR}}_{avg})/10}, \qquad (7)$$

where m is the data length and $\widehat{\mathrm{SNR}}_{avg}$ is the estimated average SNR defined as

$$\widehat{\mathrm{SNR}}_{avg} = 10\log_{10}\frac{\frac{1}{\hat{n}}\sum_{i=1}^{\hat{n}}\widehat{\alpha}_i^2}{\widehat{\sigma^2}}. \qquad (8)$$

Alternatively, we can use the MDL information criterion to eliminate the noise modes. The MDL information criterion measures the complexity of the model. Given a set of estimates and the corresponding MDL, we first discard a possible noise mode which, in general, is the smallest energy mode. We then refine the estimates using the ML procedure and compute the corresponding MDL. We compare the latter computed MDL to the given MDL. If the latter MDL is smaller than the previous case. We use the second set of estimates as the new estimates. We repeat the procedure until the best model structure which has the smallest MDL is found. The MDL information criterion for our case is found to be

$$\mathrm{MDL} = m\log\left(L(\widehat{\theta})\right) + \frac{3\widehat{n}}{2}\log(m), \qquad (9)$$

where $L(\widehat{\theta})$, m, and $\widehat{n}$ are defined as above. Note that with $m = 25$, the above MDL criterion is equivalent to a χ^2 test with a significance about 0.02 (Söderström and Stoica, 1989).

3.4. *ALGORITHM SUMMARY*

Below we summarize the model-order-revealing ML algorithm presented above:

1. Obtain the locations, magnitudes, and phases of the FFT peaks of the signal. Obtain the estimated noise power $\widehat{\sigma^2}$.
2. Calculate the resolution limit and the maximum number of modes per Fbin. Split the peak locations according to the splitting algorithm described above.
3. Perform ML estimation on the split estimates using nonlinear minimization of the negative log-likelihood function (Equation 4); using an algorithm in *e.g.*, (Ying *et al.*, 1993; Stoica *et al.*, 1989; Ziskind and Wax, 1988).
4. Check frequency separations of the estimates. Discard the small energy estimates which are within the resolution limit of some larger energy mode. Also, discard the possible noise estimates using the energy criterion defined in Equation 7 or the MDL information criterion in Equation 9.
5. If any of modes are discarded in step 4, go to Step 3. If not, the estimation is completed.

4. EXAMPLES

In this section we provide computer simulations to illustrate the performance of the proposed algorithm. For the following examples, a windowed FFT is used. In particular, a Kaiser window with

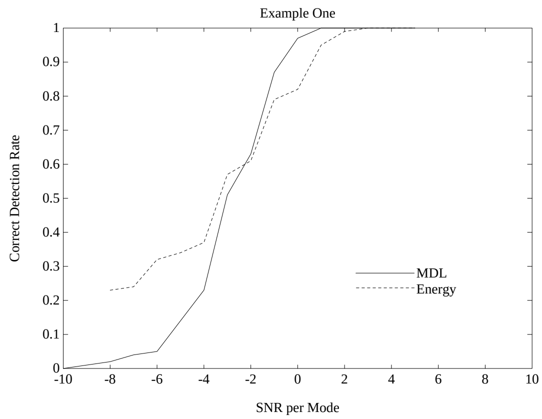

Fig. 2. Simulations results of Example One using the proposed model order determination algorithm

$\beta = 2.5$ is used (it is a trade-off between a rectangular window and a Hamming window). Also, we zero pad the data from the original length to 8192 to increase the accuracy on peak locations. We set the peak threshold to be twice the mean of the FFT magnitudes.

4.1. *EXAMPLE ONE: TWO WELL SEPARATED MODES*

In the first example we use a data sequence which is composed of $n = 2$ exponentials at frequencies of $2\pi/m$ and $-2\pi/m$ where $m = 25$ is the data length. The magnitudes of the exponentials are chosen to be one. The phases are zero.

Figure 2 shows the correct model order detection rate versus SNR per mode for this example, using 100 independent Monte-Carlo simulations. From the plot, we can see that the proposed algorithm successfully detects the model order for SNR/mode above 3 dB using the energy criterion and 1 dB using the MDL criterion. The algorithm using the energy criterion still has a success rate of about 82% for SNR/mode = 0 dB, 58% for SNR/mode = -3 dB, and 33% for SNR/mode = -6 dB. For most of the cases in which the model order is miss-estimated, we found that the criterion in Equation 7 fails to reliably discard the noise modes. For the MDL information criterion, we found that most of miss-estimated cases are underestimated. This is because, at low SNRs, the penalty of more complex model is stressed in the MDL criterion. We also found a few cases in which the pruning algorithm could not detect the extra modes because the overdetermined model was trapped in a local minimum.

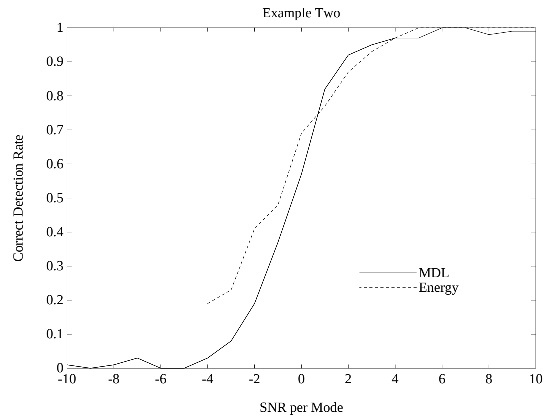

Fig. 3. Simulations results of Example Two using the proposed model order determination algorithm

4.2. *EXAMPLE TWO: TWO CLOSELY SPACED MODES*

In this example we use the following data. For $q = 0, 1, \ldots, 24$,

$$x_q = e^{j(2\pi(0.52)q+\pi/4)} + e^{j2\pi(0.5)q}.$$

This is the Kumaresan and Tufts example (Kumaresan and Tufts, 1982; Rahman and Yu, 1987).

Figure 3 also shows the correct detection rate versus SNR per mode. Again, 100 independent Monte-Carlo simulations were run for the statistics. From the figure, we see that the algorithm using both criteria can correctly detect the model order for SNR/mode above 5 dB with the fact that MDL has larger false alarm. Note that for this case (frequency separation 0.5 Fbin) the CRB resolution limit in Figure 1 is about 4 dB.

In this example we found that the algorithm tends to underestimate the model order below the critical SNR for both the energy and the MDL criterion. This is because the ML algorithm cannot resolve the two close modes, and the pruning algorithm eliminates all but this one combined mode.

5. CONCLUSIONS

We have presented an approach for estimating the model order of complex exponential signals. The detection and estimation (of the model parameters) problem were treated simultaneously in the proposed algorithm.

The statistical properties of the proposed algorithm were discussed. The algorithm first overestimates the model order as a result of the splitting procedure and then prunes the extra modes using the pruning algorithm. The splitting and pruning algorithm were developed based on the resolution limit which is derived from the CRB.

Computer simulations were provided to demonstrate the performance of the proposed algorithm. We have shown that for well separated modes the proposed algorithm can detect correctly down to 3 dB using the energy criterion and 1 dB using the MDL criterion. Under the condition of two closely spaced modes, our approach can accurately detect the model order down to about 5 dB for both criteria.

6. REFERENCES

Bakamidis, S., M. Dendrinos and G. Carayannis (1991). SVD analysis by synthesis of harmonic signals. *IEEE Transactions on Signal Processing* **39**(2), 472–477.

Bresler, Y. and A. Macovski (1986). Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP–34**(5), 1361–1375.

Cadzow, J. A. (1982). Spectral estimation: A overdetermined retional model equation approach. In: *Proceedings of the IEEE*. Vol. 70. pp. 907–939.

Harris, Fredric J. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE* pp. 51–83.

Kumaresan, R. and D. W. Tufts (1982). Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-30**(6), 833–840.

Lee, D., R. Kashyap and R. Madan (1990). Robust decentralized direction-of-arrival estimation in contaminated noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-38**(3), 496–505.

Lee, H. B. (1992). The Crameŕ-Rao Bound on Frequency Estimates of Signals Closely Spaced in Frequency. *IEEE Transactions on Signal Processing* **40**(6), 1508–1517.

Parthasarathy, S. and D. Tufts (1987). Excitation-synchronous modeling of voiced speech. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-35**(9), 1241–1249.

Rahman, M. A. and K.-B. Yu (1987). Total least squares approach for frequency estimation using linear prediction. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP–35**(10), 1440–1454.

Rao, S. and D. Gnanaprakasam (1988). A criterion for identifying dominant singular values in the SVD based method of harmonic retrieval. In: *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*. pp. 2460–2463.

Söderström, T. and P. Stoica (1989). *System Identification*. Prentice-Hall. London.

Steedly, W. and R. Moses (1993). The Cramér-Rao bound for pole and amplitude estimates of damped exponential signals in noise. *IEEE Transactions on Signal Processing* **41**(2), 1305–1318.

Steedly, W. M. and R. L. Moses (1991). High resolution exponential modeling of fully polarized radar returns. *IEEE Transactions on Aerospace and Electronic Systems* **AES-27**(3), 459–469.

Stoica, P., R. Moses, B. Friedlander and T. Söderdtröm (1989). Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-37**(3), 378–392.

Stoica, P., T. Söderström and V. Šimonytė (1992). On Estimating the Noise Power in Array Processing. *Signal Processing* **26**, 205–220.

Wax, M. and I. Ziskind (1989). Detection of the number of coherent signals by the MDL principle. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-37**(8), 1190–1196.

Wax, M. and T. Kailath (1985). Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-33**, 387–392.

Wu, Q. and D. Fuhrmann (1991). A parametric method for determining the number of signals in narrow-band direction finding. *IEEE Transactions on Signal Processing* **SP-39**(8), 1848–1857.

Yin, Y. and P. Krishnaiah (1987). On some non-parametric methods for detection of the number of signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-35**(11), 1533–1538.

Ying, C. J., S. C. Ahalt and R. L. Moses (1993). Maximum Likelihood Estimation of Exponential Signals Using Artificial Neural Networks. In: *in Proceedings of WCNN 93*. Portland, Oregon. pp. VI–629–VI–632.

Ziskind, I. and M. Wax (1988). Maximum likelihood localization of multiple sources by alternating projection. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP–36**(10), 1553–1560.