INNOVARE JOURNAL OF ENGINEERING & TECHNOLOGY

Vol 6, Issue 1, 2018

ISSN - 2347-1573

Review Article

EVALUATION OF MAXIMUM ENTROPY METHOD OF SPECTRUM ESTIMATION

JAWAHAR A[1], MURALI KRISHNA P[2], KIRAN SS

[1]Department of EEE, Sanketika Institute of Technology and Management, Visakhapatnam, Andhra Pradesh, India. [2]Department Electrical Engineer, National Operation and Maintenance Company Limited, Jeddah, Saudi Arabia. [3]Department of ECE, Sanketika Institute of Technology and Management, Visakhapatnam, Andhra Pradesh, India. Email: sskiran88k@gmail.com

Received: 05 March 2018, Revised and Accepted: 03 May 2018

ABSTRACT

The parametric models autoregressive (AR)/AR-moving average (MA)/MA are sometimes not capable of finding out the power spectral densities of random sequences. Under such circumstances, the non-parametric methods outperform the parametric ones because of the sensitivity of the latter to model specifications. The maximum entropy method (MEM) is regarded as the non-parametric method of spectrum estimation; it suggests one possible way of extrapolating the autocorrelation sequence so that a more accurate estimate of the spectrum can be obtained with better resolution. This paper investigates the work of realizing MEM method and evaluating its performance with minimum variance method.

Keywords: Minimum variance method, Maximum entropy method, Random process.

INTRODUCTION

The spectrum estimation considers the problem of estimating the power spectral density of a wide sense stationary (WSS) random process using statistical descriptors [1]. The non-parametric methods are defined as the methods of spectrum estimation which are based on the idea of estimating the autocorrelation sequence of a random process from a set of measured data, and then taking the Fourier transform to obtain an estimate of the power spectrum. These methods are classified as classical and non-classical methods. The periodogram, modified periodogram, Bartlett, Welch, Blackman-Tukey methods come under the classification of classical methods whereas the minimum variance (MV) and the maximum entropy methods (MEM) come under the classification of non-classical methods of spectrum estimation.

The classical methods effectively extrapolate the autocorrelation sequence with zeros, but one of the limitations with this approach to spectrum estimation is that the autocorrelation sequence can only be estimated for lags less than the data record length. As a result, the autocorrelation sequence is set to zero for lags greater than or equal to the data record length. This windowing significantly limits the resolution and accuracy of the estimated spectrum [2-6].

However, many random signals of interest have autocorrelations that are nonzero for lags greater than or equal to the data record length, so a more accurate extrapolation of the autocorrelation sequence is needed to mitigate the effects of the window and to find a more accurate estimate of the spectrum. The MEM method suggests one possible way to perform this extrapolation. A brief description of the non-classical methods of spectrum estimation is given in the sections 1.1 and 1.2.

THE MV METHOD

The MV method of spectrum estimation is an adaptation of the maximum likelihood method developed by Capon for the analysis of two-dimensional power spectral densities. In the MV method, the power spectrum is estimated by filtering a process with a bank of narrowband bandpass filters. The MV spectrum estimation technique involves the following steps:

1. Design a bank of bandpass filters $g_i(n)$ with a center frequency ω_i so that each filter rejects the maximum amount of out-of-band power while passing the component at frequency ω_i with no distortion.
2. Filter $x(n)$ with each filter in the filter bank and estimate the power in each output process $y_i(n)$.
3. Set $\hat{p}_x(e^{jwi})$ equal to the power estimated in step (2) divided by the filter bandwidth.

For a wide-sense stationary process $x(n)$, the MV spectrum estimate of the power spectrum is

$$\hat{P}_{MV}(e^{jw}) = \frac{p+1}{e^H R_x^{-1} e} \quad (1)$$

Where R_x is the $p \times p$ autocorrelation matrix. The resolution of the spectrum estimate of the MV method is less, hence, we go for MEM method of spectrum estimation [1].

MEM

An important application of entropy (uncertainty) is the determination of the probabilities of the random events of a partition subject to various constraints, with the MEM. The method of MEM states that the unknown probabilities must be so chosen as to maximize the entropy (maximize the uncertainty) of the partition subject to the given constraints [2]. It is optimal for problems where we have prior information about multiplicities, but no noise [3]. Its estimate is based on the principle that the estimate of the autocorrelation sequence must correspond to the most random signal whose autocorrelation values in the lag values less than or equal to the finite range coincide with the measured values [4]. The ME method finds an all-pole model for a process using the autocorrelation method and then uses the model parameters to estimate the spectrum [1]. It is represented as follows:

$$\hat{P}_{mem}(e^{jw}) = \frac{\in p}{\left|1 + \sum_{k=1}^{p} a_p(k) e^{-jkw}\right|^2} \quad (2)$$

Where $a_{p(K_j)}$ and ϵ_p are the all-pole coefficients.

MATHEMATICAL MODELING OF MEM

Given the autocorrelation $r_x(\kappa)$ of a WSS process for lags $|\kappa| \geq p$, the problem that we wish to address, illustrated in Fig. 1, is how to extrapolate $r_x(\kappa)$ for $|\kappa| > p$. Denoting the extrapolated values by $r_e(\kappa)$ it is clear that some constraints should be placed on $r_e(\kappa)$.

For example, if

$$P_x(e^{jw}) = \sum_{k=-p}^{p} r_x(k)e^{-jkw} + \sum_{|k|>p} r_e(k)e^{-jkw} \quad (3)$$

Then, $P_x(e^{ji})$ should correspond to a valid power spectrum, i.e. $P_x(e^{j\omega})$ should be real-valued and non-negative for all ω. In general, however, only constraining $P_x(e^{j\omega})$ to be real and non-negative is not sufficient to guarantee a unique extrapolation. Therefore, some additional constraints must be imposed on the set of allowable extrapolations. One such constraint proposed by Burg is to perform the extrapolation in such a way so as to maximize the entropy of the process. Since entropy is a measure of randomness or uncertainty, a maximum entropy extrapolation is equivalent to finding the sequence of autocorrelations, $r_e(\kappa)$, that make $_x(n)$ as white (random) as possible. In some sense, such an extrapolation places as few constraints as possible or the least amount of structure on $_{x(n)}$. In terms of the power spectrum, this corresponds to the constraint that $P_x(e^{j\omega})$ be "as flat as possible" (Fig. 2).

For a Gaussian random process with power spectrum $P_x(e^{j\omega})$, the entropy is

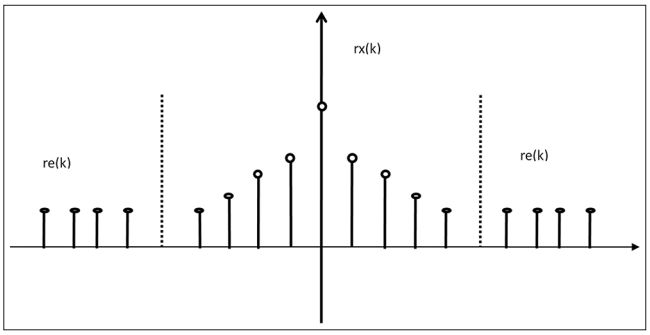

Fig. 1: Extrapolating autocorrelation sequence

$$H(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln P_X(e^{jw})d \quad (4)$$

Therefore, for Gaussian processes with a given partial autocorrelation sequence, $r_x(\kappa)$ for $|\kappa| \leq p$, the maximum entropy power spectrum is the one that maximizes Eq. (4) subject to the constraint that the inverse discrete-time Fourier transform of $P_x(e^{j\omega})$ equals the given set of autocorrelations for $|\kappa| \leq p$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_X(e^{jw})e^{jkw}dw = r_x(k); |k| \leq p \qquad |k| \leq p \quad (5)$$

The values of $r_e(\kappa)$ that maximize the entropy may be found by setting the derivative of $H(x)$ with respect to equal to zero as follows:

$$\frac{\partial H(x)}{\partial r_e^*(k)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{P_X(e^{jw})}\frac{\partial P_X(e^{jw})}{\partial r_e^*(k)}dw = 0; \quad |k| > p \quad (6)$$

From Eq. (1) we see that

$$\frac{\partial P_X(e^{jw})}{\partial r_e^*(k)} = e^{jkw} \quad (7)$$

Which, when substituted into Eq. (6), yields

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{P_X(e^{jw})}e^{jkw}dw = 0; \quad |k| > p \quad (8)$$

Defining $Q_x(e^{j\omega})=1/P_x(e^{jk\omega}_{})$, Eq. (8) states that the inverse discrete-time Fourier transforms of $Q_x(e^{j\omega})$ is a finite-length sequence that is equal to zero for $|\kappa| > p$.

$$q_X(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} Q_X(e^{jw})e^{jkw}dw = 0; \quad |k| > p \quad (9)$$

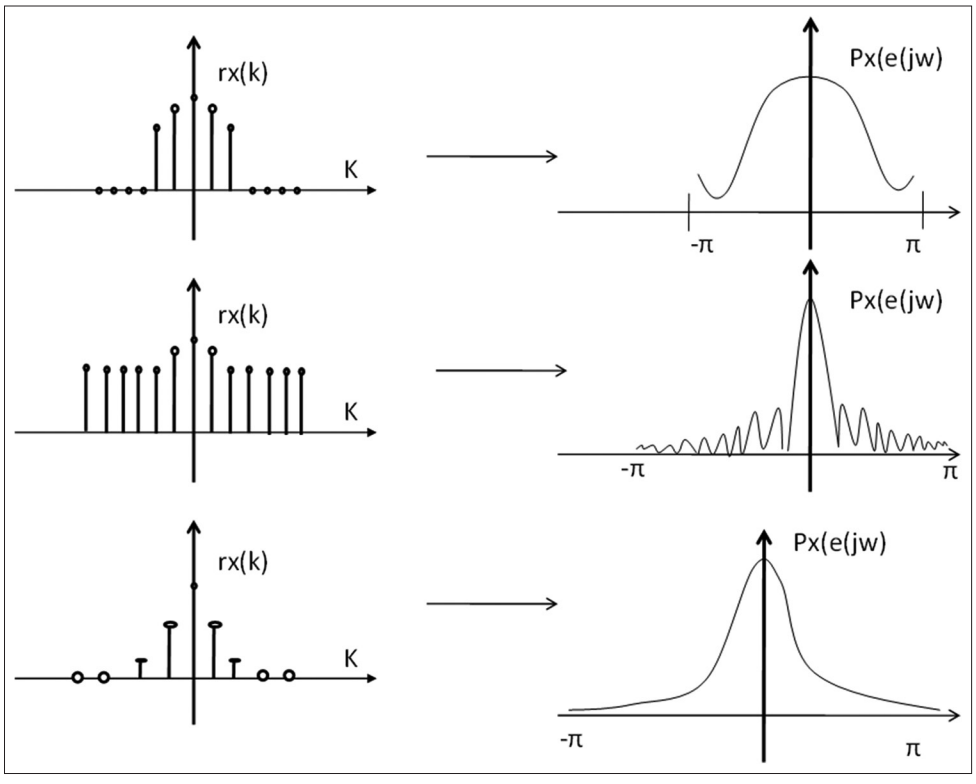

Fig. 2: Different extrapolations of a partial autocorrelation sequence and the corresponding power spectral densities

Therefore,

$$Q_X(e^{jw}) = \frac{1}{P_X(e^{jw})} = \sum_{k=-p}^{p} q_x(k)e^{-jkw} \tag{10}$$

it follows that the minimum entropy power spectrum for a Gaussian process, which we will denote by $Pmem(e^{j\omega})$, is an all-pole power spectrum,

$$\hat{P}_{mem}(e^{jw}) = \frac{1}{\sum_{k=-p}^{p} q_X(k)e^{-jkw}} \tag{11}$$

Using the spectral factorization theorem, it follows that Eq. (11) may be expressed as

$$\hat{P}_{mem}(e^{jw}) = \frac{|b(0)|^2}{A_p(e^{jw})A_p^*(e^{jw})} = \frac{|b(0)|^2}{|1+\sum_{k=-p}^{p} a_p(k)e^{-jkw}|^2} \tag{12}$$

Alternatively, in terms of the vectors

$\alpha_p=[1,\alpha p(1),\ldots,\alpha p(p)]^T$ and $e=[1,e^{j\omega},\ldots,e^{jp\omega}]T$, the MEM spectrum may be written as

$$\hat{P}_{mem}(e^{jw}) = \frac{|b(0)|^2}{|e^H a_p|^2} \tag{13}$$

Having determined the form of the MEM spectrum, all that remains is to find the coefficients $\alpha_p(\kappa)$ and $b(0)$. Due to the constraint is given in Eq. (5), these coefficients must be chosen in such a way that the inverse discrete-time Fourier transform of $Pmem(e^{j\omega})$ produces an autocorrelation sequence that matches the given values of $r_x(\kappa)$ for $|\kappa|\leq p$. If the coefficients $\alpha_p(\kappa)$ are the solution to the autocorrelation normal equations:

$$\begin{bmatrix} r_X(0) & r_X^*(1) & \cdots & r_X^*(p) \\ r_X(1) & r_X(0) & \cdots & r_X^*(p-1) \\ \vdots & \vdots & & \vdots \\ r_X(p) & r_X(p-1) & \cdots & r_X(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_p \\ \vdots \\ a_p(p) \end{bmatrix} = \in_p \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{14}$$

And if:

$$|b(0)|^2 = r_X(0) + \sum_{k=1}^{p} a_p(k) r_X^*(k) = \in_p \tag{15}$$

Then, the autocorrelation matching constraint given in Eq. (5) will be satisfied. Thus, the MEM spectrum is:

$$\hat{P}_{mem}(e^{jw}) = \frac{\in_p}{\left|e^H a_p\right|^2} \tag{16}$$

Where a_p is the solution to Eq. (14). In summary, given a sequence of autocorrelations, $r_x(\kappa)$ for $\kappa=0,1,\ldots,p$, the MEM spectrum is computed as follows. First, the autocorrelation normal equations (14) are solved for the all-pole coefficients $\alpha_p(\kappa)$ and ϵ_p. Then, the MEM spectrum is formed by incorporating these parameters into Eq. (16). Note that since $\hat{P}mem(e^{j\omega})$ is an all pole power spectrum, then $r_x(\kappa)$ satisfies the Yule-Walker equations.

$$r_X(l) = -\sum_{k=1}^{p} a_p(k) r_X(k-l) \text{ for } l>0 \tag{17}$$

Therefore, the MEM extrapolates the autocorrelation sequence according to this recursion. The properties of MEM method have been studied extensively and as a spectrum analysis tool, this method is subject to different interpretations. It may be argued, for example, that in the absence of any information or constraints on a process $x(n)$, given a set of autocorrelation values, $r_x(0),\ldots rx(p)$, the best way to estimate the power spectrum is to Fourier transform the autocorrelation sequence formed from the given values along with an extrapolation that imposes the least amount of structure on the data, i.e., performs a maximum entropy extrapolation. This would seem to be preferable of an extrapolation that somewhat arbitrarily sets $r_x(\kappa)=0$ for $|\kappa|>p$ as in the classical approach. On the other hand, it may also be argued that since the maximum entropy extrapolation imposes an all-pole model on the data unless the process is known to be consistent with this model, then the estimated spectrum may not be very accurate. The maximum entropy spectrum is identical to the autoregressive (AR)-model spectrum only when the exact autocorrelation $r_x(\kappa)$ is known. When only an estimate of $r_x(\kappa)$ is available for $0\leq\kappa\leq p$, the AR-model estimates of Yule-walker and Burg are not maximum entropy spectrum based on estimates of the autocorrelation sequence results in a set of nonlinear equations [5].

RELATIONSHIP BETWEEN MEM AND MV SPECTRUM ESTIMATES

There is an interesting relationship that exists between the MEM and the MV spectrum estimates. This relationship states that the p th-order MV estimate is the harmonic mean of the MEM estimates of orders $\kappa=0,1,\ldots,p$. To derive this relationship, we will use the recursion given in Eq. below

$$R_{jH}^{-1} = \begin{bmatrix} 0 & 0^T \\ 0 & R_j^{-1} \end{bmatrix} + \frac{1}{\in_{j+1}} a_{j+1} a_{j+1}^H \tag{18}$$

for computing the inverse of a Toeplitz matrix. This recursion provides the following expression for the inverse of the (p+1)×(p+1) Toeplitz matrix R_p,

$$R_p^{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & R_{p-1}^{-1} & \\ 0 & & & \end{bmatrix} + \frac{1}{\varepsilon_p} a_p a_p^H \tag{19}$$

For a WSS process with an autocorrelation matrix R_p, the p th-order MV estimate is

$$P_{MV}(e^{jw}) = \frac{p+1}{e^H R_p^{-1} e} \tag{20}$$

Multiplying on the left by e^H and on the right by e we have

$$e^H R_p^{-1} e = e^H \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & R_{p-1}^{-1} & \\ 0 & & & \end{bmatrix} e + \frac{1}{\varepsilon_p} e^H a_p a_p^H e \tag{21}$$

or,

$$e^H R_p^{-1} e = e^H R_{p-1}^{-1} e + \frac{1}{\in_p} \left|e^H a_p\right|^2 \tag{22}$$

Note that the first two terms in Eq. (22) are proportional to the inverse of the MV estimates of order p and $p-1$, respectively, and the last term is the reciprocal of the p th-order MEM spectrum. Therefore, we have

$$\frac{p+1}{\hat{P}_{MV}^{(p)}(e^{jw})} = \frac{p}{\hat{P}_{MV}^{(p-1)}(e^{jw})} + \frac{1}{\hat{P}_{mem}^{(p)}(e^{jw})} \tag{23}$$

which is a recursion for the p th-order MV estimate in terms of the $(p-1)$ st-order MV estimate and the p th-order MEM spectrum. Solving this recursion for we find

$$\frac{1}{\hat{P}_{MV}^{(p)}(e^{jw})} = \frac{1}{p+1} \sum_{k=0}^{p} \frac{1}{\hat{P}_{mem}^{(k)}(e^{jw})} \tag{24}$$

Therefore, the MV estimate is the harmonic mean of the MEM spectra from the low order to the high order estimates. As a result of this

smoothing, for WSS processes consisting of narrowband components in noise, the MEM spectrum generally provides a higher resolution spectrum estimate than the MV method.

SELECTION CRITERIA FOR PERFORMANCE EVALUATION

An important factor in the selection of a spectrum estimation technique is the performance of the estimator. In comparing one non-parametric method to another, there is a trade-off between resolution and variance. The variability, v of the estimate is represented as,

$$v = \frac{\mathrm{var}\left\{\hat{P}_x(e^{jw})\right\}}{E^2\left\{\hat{P}_x(ew)\right\}} \quad (25)$$

The variability must be as low as possible to determine the given non-parametric method as the best method.

Resolution, Δw of the estimate is represented as,

$$\Delta w = f_2 - f_1 \quad (26)$$

Where $f_2 - f_1$ is the bandwidth of the main lobe.

The resolution must be high to determine the given non-parametric method as the best method.

The overall figure of merit μ is defined as the product of the variability, v and the resolution Δw.

$$\mu = v\Delta w \quad (27)$$

As the figure of merit decreases the performance of the non-parametric method increases, so the figure of merit should be as low as possible.

MONTE CARLO SIMULATION AND RESULTS

For the purpose of simulation, the following random process consisting of a sinusoidal signal in white noise is considered.

$$x(n) = 5\sin(n\omega) + v(n) \quad (28)$$

With $\omega = 0.4\pi$, $n = 512$ and (n) is zero mean unit variance white noise.

The order of the bandpass filter (p) is taken as 10.

Realization of MV method

Monte Carlo simulation with 50 runs of the input signal shown in Eq. (28) is performed to evaluate the performance of the MV spectrum. The steps involved in this evaluation are as follows:

Step 1: The signal shown in Eq. (28) is transposed.
Step 2: Covariance matrix of the transposed input signal and the bandpass filter is computed.
Step 3: Eigenvalues and the corresponding eigenvectors are computed so that the criterion R*v=v*d is satisfied.
Step 6: A diagonal matrix which is the inverse of the absolute of the eigenvalues is computed.
Step 7: The squared value of the power spectrum of the eigenvectors is calculated.
Step 8: Frequency is normalized for the purpose of plotting the MV spectrum.
Step 9: The periodogram of the MV method is computed and plotted in Fig. 3.

Realization of MEM

Monte Carlo simulation with 50 runs of the input signal shown in Eq. (28) is performed to evaluate the performance of the MEM spectrum. The steps involved in this evaluation are as follows:

Step 1: The random process consisting of noise shown in Eq. (28) is transposed.
Step 2: Frequency is normalized for the purpose of plotting. The all-pole coefficients α and e are calculated using the autocorrelation matrix.
Step 3: The inputs of autocorrelation matrix are x' and p.
Step 4: The MEM spectrum is computed using Eq. (16) and plotted in Fig. 4.

Comparison of MEM and MV methods

The type of process being analyzed and the closeness of the process to the AR process are the critical factors on which the comparison of the MEM with MV method depends.

Table 1: Comparison of non-classical methods using simulated results

Method used	Variability	Resolution	Figure of merit
MV	0.0256	0.15	0.0038
Maximum entropy	0.0055	0.3	0.0016

MV: Minimum variance

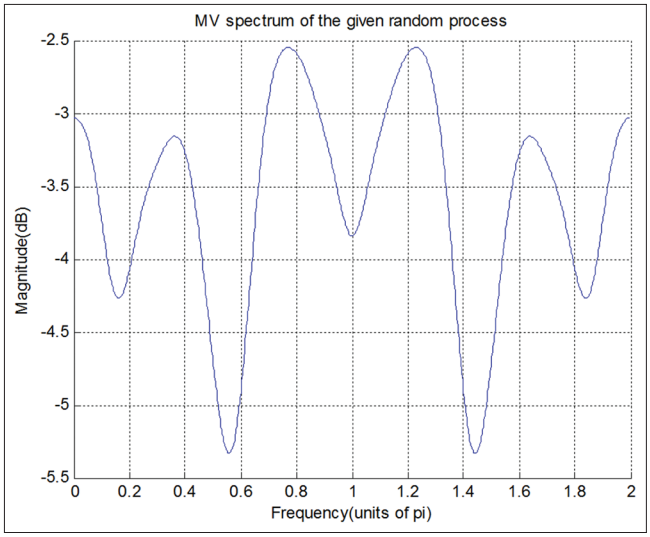

Fig. 3: Minimum variance spectrum of the given random process

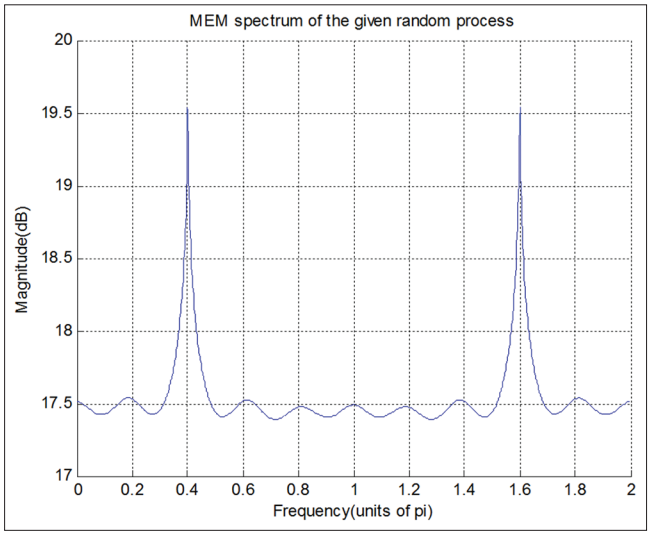

Fig. 4: Maximum entropy method spectrum of the given random process

CONCLUSION

The simulated results as shown in Table 1, in section 5.3, show that the performance of the MEM is better than that of the MV method as it produced less variability, the figure of merit, and high resolution than MV method.

REFERENCES

1. Monson H. Hayes, Statiscal Digital Signal Processing and Modeling. New York: John Wiley and Sons, INC; 2002.
2. Papoulis A. Probability, Random Variables, and Stochastic Processes. 3rd ed. New York, NY: McGraw-Hill Europe; 2002.
3. Janes ET. On the Rationale of Max-Ent Method. IEEE Paper
4. Vaseghi SV. Advanced Digital Signal Processing and Noise Reduction. 3rd ed. Chichester: John Wiley & Sons, INC; 2006.
5. Proakis JG, Manolakis DG. Digital Signal Processing Principles, Algorithms and Applications, PHI: Prentice. Hall; 2007.
6. Chapman SJ. "MATLAB Programming for Engineers. 3rd ed. Pacific Grove, CA: Brooks/Cole; 2004.