Development and Analysis of a Neural Network Approach to Pisarenko's Harmonic Retrieval Method

George Mathew and V. U. Reddy

Abstract—Pisarenko's harmonic retrieval (PHR) method is perhaps the first eigenstructure based spectral estimation technique. The basic step in this method is the computation of eigenvector corresponding to the minimum eigenvalue of the autocorrelation matrix of the underlying data. In this paper, we recast a known constrained minimization formulation for obtaining this eigenvector into the neural network (NN) framework. Using the penalty function approach, we develop an appropriate energy function for the NN. This NN is of feedback type with the neurons having sigmoidal activation function. Analysis of the proposed approach shows that the required eigenvector is a minimizer (with a given norm) of this energy function. Further, all its minimizers are global minimizers. Bounds on the integration time step that is required to numerically solve the system of nonlinear differential equations, which define the network dynamics, have been derived. Results of computer simulations are presented to support our analysis.

I. INTRODUCTION

ESTIMATION of the frequencies of sinusoids corrupted with white noise arises in many applications. The various spectral estimation techniques which can be applied to solve this problem can be classified into two categories; eigenstructure based methods (which depend on the eigenstructure of the covariance matrix of the underlying data) and noneigenstructure based methods. Pisarenko's harmonic retrieval (PHR) method and Maximum entropy method are examples, respectively, of these two classes. The eigenstructure based methods are preferred to the other, since they yield high resolution and asymptotically exact results. In this paper, we concentrate on the PHR method and solve the basic step in this method, i.e., estimation of the eigenvector corresponding to the minimum eigenvalue of the covariance matrix, by exploiting the optimization property of feedback type neural networks [1].

Let

$$y(n) = \sum_{i=1}^{P} \alpha_i \cos(\omega_i n + \theta_i) + v(n) \qquad (1.1)$$

where α_i, ω_i and θ_i denote the amplitude, frequency (normalized) and initial phase (uniform in $[0, 2\pi]$) of the ith sinusoid and $\{v(n)\}$ are zero mean, *iid* random variables with variance σ^2. Let $\mathbf{R}$ denote the covariance matrix of size $N \times N, (N \geq 2P + 1)$ of $y(n)$. Then, the eigenvector corresponding to the minimum eigenvalue, λ_{min}, (hereafter

referred to as the minimum eigenvector) of $\mathbf{R}$ is the solution of the following constrained minimization problem [2]

$$\min_{\mathbf{w}} \mathbf{w}^T \mathbf{R} \mathbf{w} \qquad \text{subject to} \qquad \mathbf{w}^T \mathbf{w} = 1 \qquad (1.2)$$

where $\mathbf{w} = [w_1, w_2, \cdots, w_N]^T$ is an N-dimensional weight vector. Then, the polynomial formed using the elements of this eigenvector will have $2P$ of its $N - 1$ roots located at $\exp(\pm j\omega_i), i = 1, \cdots, P$. These $2P$ roots of interest will be unaffected by the noise power and the remaining $N - 1 - 2P$ roots are arbitrary [3].

Thus, the central problem in the PHR method is the computation of a minimum eigenvector. Different techniques have been proposed for efficient and adaptive estimation of the minimum eigenvector [2], [4], [6]. While Thompson [2] suggested a constrained gradient search procedure, Reddy et al. [4] restated this constrained minimization problem as an uncontrained minimization problem and developed a Newton type recursive algorithm for seeking the minimum eigenvector. Larimore [5] studied the convergence behavior of Thompson's [2] adaptive algorithm. Using rotational search method, Fuhrmann and Liu [6] proposed two adaptive algorithms for the adaptive PHR method.

Many neural network (NN) based algorithms have been reported [8]–[12] for eigenvector estimation. All these algorithms are developed for feedforward NN structure. They estimate the principal eigenvectors of the covariance matrix of the sequence of input vectors. Oja [8] proposed a Hebbian type learning rule for a single neuron model and showed its convergence to the first principal eigenvector. The problem of learning in a two (or more) layered NN by minimizing a quadratic error function was considered by Baldi and Hornik [9]. They proved that the error function has a unique minimum corresponding to the projection onto the subspace generated by the principal eigenvectors and all other stationary points are saddle points. Sanger [10] proposed the generalized Hebbian algorithm for a single layer NN and showed its convergence to the principal eigenvectors. Kung and Diamantaras [11] proposed an approach for the recursive computation of principal eigenvectors. Their technique combines the Hebbian and orthogonal learning rules. Kung [12] extended this approach to the case of constrained principal component analysis.

In this paper, we suggest a NN approach to the PHR problem. Note that while all the above mentioned NN algorithms seek principal eigenvectors, we seek the minimum eigenvector. Instead of feedforward networks with linear neurons, we use feedback network with sigmoidal neurons. Our approach is purely an optimization based one and the cost function

used is different from that of others except Chauvin [13] who used a similar cost function for estimating the principal eigenvector with a linear single neuron model. Our approach is quite general in that it can be applied to estimate the minimum eigenvector of a general symmetric matrix which is not indefinite. But, none of the above algorithms have this feature.

This paper is organized as follows. The NN formulation of the PHR problem is presented in Section II. Analysis of the proposed approach is presented in Section III. This includes description of the landscape of the energy function and convergence of the network dynamics. Simulation results are presented in Section IV and Section V concludes the paper.

II. NEURAL NETWORK FORMULATION OF THE PHR PROBLEM

From the theory of feedback neural networks, we know that a system of N neurons connected in feedback configuration evolves in such a way that the stable stationary states of the network correspond to the minima of the energy function (or Lyapunov function) of the network. Hence, we need to construct an appropriate energy function whose minima would correspond to the minimum eigenvectors of $\mathbf{R}$.

The cost function used in our development is motivated as follows. Using the penalty function method [7], (1.2) can be replaced by an unconstrained problem of the form

$$\min_{\mathbf{w}} \quad \{J(\mathbf{w}, \mu) = \mathbf{w}^T \mathbf{R} \mathbf{w} + \mu P(\mathbf{w})\}$$
$$P(\mathbf{w}) = (\mathbf{w}^T \mathbf{w} - 1)^2 \qquad (2.1)$$

where μ is a positive constant and P is a penalty function satisfying three conditions: (1) P is continuous, (2) $P(\mathbf{w}) \geq 0 \, \forall \, \mathbf{w}$, and (3) $P(\mathbf{w}) = 0$ if and only if $\mathbf{w}^T \mathbf{w} = 1$.

Let $\{\mu_k\}, k = 1, 2, 3, \cdots$, be a sequence tending to infinity such that for each $k, \mu_k \geq 0$ and $\mu_{k+1} > \mu_k$. Further, let $\mathbf{w}_k$ be the minimizer of $J(\mathbf{w}, \mu_k)$. Then, we have the following theorem [7].

Theorem: Limit point of the sequence $\{\mathbf{w}_k\}$ is a solution of (1.2).

That is, limit point of $\{\mathbf{w}_k\}$ is a minimum eigenvector of $\mathbf{R}$ with unit norm. However, there is a special structure in the cost function (2.1), resulting in the fact (as proved in Section III) that "a minimizer of $J(\mathbf{w}, \mu)$ for any given μ (with $\mu > (\lambda_{\min}/2)$) is a minimum eigenvector of $\mathbf{R}$." Further, a minimum eigenvector with any norm would suffice for the PHR problem. Hence, it is enough to choose a μ satisfying $\mu > (\lambda_{\min}/2)$.

Our aim is to design a neural network for which J could be the energy function. This network must have N neurons with the output of kth neuron, w_k, representing the kth element of the vector $\mathbf{w}$. Sigmoidal activation function is assumed for each neuron, i.e.

$$w_k(t) = f(u_k(t)) = \frac{2}{1 + \exp{(-u_k(t))}} - 1 \qquad (2.2)$$

where $u_k(t)$ and $w_k(t)$ are the input and output, respectively, of the kth neuron at time t. For J to be the energy function, the network dynamics should be such that the time derivative of

J is negative. For the cost function (2.1), the time derivative is given by

$$\frac{dJ}{dt} = \sum_{k=1}^{N} \frac{\partial J}{\partial w_k(t)} \cdot f'(u_k(t)) \cdot \frac{du_k(t)}{dt} \qquad (2.3)$$

where $f'(u)$ is the derivative of $f(u)$ with respect to u.

Now, suppose we define the dynamics of the kth neuron as

$$\frac{du_k(t)}{dt} = -\frac{\partial J}{\partial w_k(t)}$$
$$w_k(t) = f(u_k(t)) \qquad k = 1, \cdots, N. \qquad (2.4)$$

Since $f(u)$ is a monotonically increasing function, it can be easily deduced from (2.3) and (2.4) that the NN with dynamics given by (2.4) has stable stationary points at the local minima of J. In the next section, we show that a minimizer of J corresponds to a minimum eigenvector of $\mathbf{R}$.

III. ANALYSIS OF THE PROPOSED APPROACH

In this section, we analyze the proposed neural network approach and establish it as a minimum eigenvector estimator. We do this in two steps. First, we establish the correspondence between the minimizers of J and the minimum eigenvectors of $\mathbf{R}$. This derivation also results in establishing guidelines for the selection of μ. Next, we derive the bounds on the integration time-step (h) which is used in solving the system of N differential equations (2.4) numerically.

The problem in hand is an unconstrained nonlinear optimization problem. In the following analysis, we assume that μ is fixed at some appropriately chosen positive value. The gradient vector $\mathbf{g}(\mathbf{w})$ and Hessian matrix $\mathbf{H}(\mathbf{w})$ of J, respectively, are given by

$$\mathbf{g}(\mathbf{w}) = 2\mathbf{R}\mathbf{w} + 4\mu\mathbf{w}(\mathbf{w}^T\mathbf{w} - 1) \qquad (3.1)$$
$$\mathbf{H}(\mathbf{w}) = 2\mathbf{R} + 8\mu\mathbf{w}\mathbf{w}^T + 4\mu(\mathbf{w}^T\mathbf{w} - 1)\mathbf{I}_N. \qquad (3.2)$$

Since $\mathbf{R}$ is symmetric, we can express it as

$$\mathbf{R} = \sum_{i=1}^{N} \lambda_i \mathbf{e}_i \mathbf{e}_i^T \qquad (3.3)$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{2P} > \lambda_{2P+1} = \lambda_{2P+2} = \cdots = \lambda_N = \sigma^2 > 0$ are the eigenvalues of $\mathbf{R}$ in the decreasing order and $\mathbf{e}_i$ is the orthonormalized eigenvector of $\mathbf{R}$ corresponding to λ_i. The last $N - 2P$ eigenvalues and eigenvectors $(\lambda_i, \mathbf{e}_j, j = 2P + 1, \cdots, N)$ are referred to as the minimum eigenvalues and minimum eigenvectors, respectively, of $\mathbf{R}$. Substituting (3.3) into (3.2), we get

$$\mathbf{H}(\mathbf{w}) = 2\sum_{i=1}^{N} [\lambda_i - 2\mu(1 - \mathbf{w}^T\mathbf{w})]\mathbf{e}_i\mathbf{e}_i^T + 8\mu\mathbf{w}\mathbf{w}^T. \qquad (3.4)$$

A. Correspondence Between the Minimizers of J and the Minimum Eigenvectors of $\mathbf{R}$

Because of the use of sigmoidal nonlinearity, the state space of the NN (i.e., the space in which $\mathbf{w}$ lies) is the N-dimensional unit open hypercube which we denote by S.

Theorem 1: $\mathbf{w}$ is a stationary point (SP) of J if and only if $\mathbf{w}$ is an eigenvector of $\mathbf{R}$ corresponding to the eigenvalue λ, with $\|\mathbf{w}\|_2 = \beta = \sqrt{1 - \frac{\lambda}{2\mu}}$.

This result immediately follows from (3.1).

Theorem 2: $\mathbf{w}$ is a global minimizer of J if and only if $\mathbf{w}$ is a minimum eigenvector of $\mathbf{R}$ corresponding to the eigenvalue $\lambda_{\min}$, with $\|\mathbf{w}\|_2 = \beta = \sqrt{1 - (\lambda_{\min}/(2\mu))}$.

Proof:

If part: From the hypothesis, we have

$$\mathbf{R}\mathbf{w} = \lambda_{\min}\mathbf{w} \quad \text{and} \quad \lambda_{\min} = 2\mu(1 - \beta^2). \quad (3.5)$$

Hence, from Theorem 1, $\mathbf{w}$ is a SP of J. To prove that $\mathbf{w}$ is a global minimizer of J, we have to show that $J(\mathbf{x}) - J(\mathbf{w}) \geq 0 \ \forall \mathbf{x} \in S$.

Let $\mathbf{x} = \mathbf{w} + \mathbf{p}$, where, $\mathbf{p} \in S$. Then, evaluating J at $\mathbf{x}$ and simplifying, we get

$$J(\mathbf{x}) - J(\mathbf{w}) = 2\mathbf{p}^T\mathbf{R}\mathbf{w} + \mathbf{p}^T\mathbf{R}\mathbf{p} + \mu[2\mathbf{p}^T\mathbf{w} + \mathbf{p}^T\mathbf{p}]^2$$
$$+ 2\mu(\beta^2 - 1) \cdot [2\mathbf{p}^T\mathbf{w} + \mathbf{p}^T\mathbf{p}]. \quad (3.6)$$

Note that $\mathbf{p}^T\mathbf{R}\mathbf{p} \geq \lambda_{\min}\mathbf{p}^T\mathbf{p}$ for any $\mathbf{p}$. Substituting this and (3.5) into (3.6), we get

$$J(\mathbf{x}) - J(\mathbf{w}) \geq \mu[2\mathbf{p}^T\mathbf{w} + \mathbf{p}^T\mathbf{p}]^2 \geq 0 \quad \forall \mathbf{x} \in S \quad (3.7)$$

which implies that $\mathbf{w}$ is a global minimizer of J.

Only if part: From the hypothesis (and using Theorem 1), we get

$$\mathbf{R}\mathbf{W} = \lambda_m\mathbf{w} \quad \text{with} \quad \lambda_m = 2\mu(1 - \beta^2)$$
$$\text{for some} \quad m \in \{1, \cdots, N\}. \quad (3.8)$$

Substituting (3.8) in (3.4) and expressing $\mathbf{w} = \beta\mathbf{e}_m$, we obtain

$$\mathbf{H}(\mathbf{w}) = \sum_{i=1}^{N} h_i\mathbf{e}_i\mathbf{e}_i^T$$

where $h_i = 2(\lambda_i - \lambda_m) + 8\mu\beta^2\delta_{mi}$ is the ith eigenvalue of $\mathbf{H}(\mathbf{w})$ with $\mathbf{e}_i$ as the corresponding eigenvector, and δ_{mi} is the Kronecker delta function. Since $\mathbf{w}$ is a global minimizer, $\mathbf{H}(\mathbf{w})$ should be at least positive semidefinite. Hence, we get

$$2(\lambda_i - \lambda_m) + 8\mu\beta^2\delta_{mi} \geq 0 \quad \forall i = 1, \cdots, N. \quad (3.9)$$

As $\lambda_i \geq \lambda_{\min} \forall i$, (3.9) can be true only if $\lambda_m = \lambda_{\min}$. This implies that $m \in \{2P + 1, \cdots, N\}$ and thus $\mathbf{w}$ is a minimum eigenvector with $\beta^2 = (1 - \lambda_{\min}/2\mu)$.

We state four corollaries below to bring out the significant features of Theorem 2.

Corollary 1: The value of μ should be such that $\mu > (\lambda_{\min}/2)$.

Corollary 2: For a given μ, every local minimizer of J is also a global minimizer.

This result follows from Theorem 2 by recognizing that $\mathbf{H}(\mathbf{w})$ is at least positive semidefinite when $\mathbf{w}$ is a local minimizer of J.

Corollary 3: The minimizer of J is unique (except for the sign) only when $N = 2P + 1$.

Corollary 4: The eigenvectors of $\mathbf{R}$ associated with the first $2P$ eigenvalues correspond to saddle points of J.

This follows from the fact that $\mathbf{H}(\mathbf{w})$ is indefinite at the stationary point $\mathbf{w}$ if it is an eigenvector corresponding to a non-minimum eigenvalue.

Discussion: When $N > 2P + 1$, minimizers of J can be expressed in terms of the $N - 2P$ orthonormal minimum eigenvectors as $\mathbf{w} = \Sigma_{i=2P+1}^{N} \alpha_i\mathbf{e}_i, \alpha_i \in \mathcal{R}$. Hence, $\mathbf{w}^T\mathbf{w} = \beta^2 = \Sigma_{i=2P+1}^{N} \alpha_i^2$. This implies that all vectors that lie on the boundary of a hypersphere of radius β in the space of minimum eigenvectors of $\mathbf{R}$ are minimizers of J. This suggests (in view of Theorem 2 and Corollary 2) that minimizers of J form a continuum on this hypersphere.

An important point to be noted from Theorem 2 is that the norm of the solution is predetermined by the values of μ and $\lambda_{\min}$. The higher the value of μ, the closer will be this norm to unity. Further, it is required to know the value of $\lambda_{\min}$ for choosing μ. Since, $\lambda_{\min}$ will not be known *a priori*, we suggest the following practical lower bound

$$\mu > \frac{\text{Trace}(\mathbf{R})}{2N}. \quad (3.10)$$

We may add here that the proposed approach can be extended to the case of a general symmetric matrix (which is not indefinite).

B. Bounds on the Integration Time-Step

The system of differential equations which defines the NN dynamics is given by (cf. (2.4))

$$\frac{d\mathbf{u}(t)}{dt} = -2\mathbf{R}\mathbf{w}(t) - 4\mu\mathbf{w}(t)[\mathbf{w}^T(t)\mathbf{w}(t) - 1] \quad (3.11)$$

$$w_k(t) = f(u_k(t)) \quad k = 1, \cdots, N \quad (3.12)$$

where $\mathbf{u}(t) = [u_1(t), \cdots, u_N(t)]^T$. The minimizer of J corresponds to the solution of this system. To solve this numerically, we need to choose an appropriate integration time-step, say h, which guarantees convergence of the technique to the correct solution. We now present an approximate convergence analysis to obtain the bounds for h, assuming a simple time-discretization numerical technique.

For sufficiently small h, we have the following approximation

$$\frac{d\mathbf{u}(t)}{dt}\Big|_{t=nh} \approx \frac{\mathbf{u}(n+1) - \mathbf{u}(n)}{h}$$

where n is the discrete time index. Then, (3.11) can be rewritten as

$$\mathbf{u}(n+1) \approx \mathbf{u}(n) - 2h\mathbf{B}(n)\mathbf{w}(n) \quad (3.13)$$

where $\mathbf{B}(n) = \mathbf{R} + 2\mu d(n)\mathbf{I}_N, d(n) = \mathbf{w}^T(n)\mathbf{w}(n) - 1$. Since the surface of J does not have local minima (cf. Corollary 2), the discrete-time gradient descent implemented by (3.13) and (3.12) with step size h will reach a global minimizer of J provided h is sufficiently small. We can therefore assume that $\mathbf{w}(n)$ is very close to the desired solution, say $\mathbf{w}^*$, for $n > K$, where K is a large enough positive integer, and that the norm of $\mathbf{w}(n)$ remains approximately constant at $\beta^* = \|\mathbf{w}^*\|_2$.

TABLE I
ESTIMATES OF FREQUENCIES, β AND $\lambda_{\min}$ FOR DIFFERENT VALUES OF μ, N AND P.

μ	N	P	True Frequencies			Estimated Frequencies			$\hat{\beta}$	$\hat{\lambda}_{\min}$
			f_1	f_2	f_3	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$		
1	5	1	0.20			0.2000			0.707356	1.00
40	5	1	0.20			0.2003			0.993730	1.00
1	5	2	0.20	0.24		0.2003	0.2402		0.707106	1.00
40	5	2	0.20	0.24		0.2002	0.2402		0.993730	1.00
1	7	1	0.20			0.2000			0.707355	1.00
40	7	1	0.20			0.2000			0.993730	1.00
1	7	2	0.20	0.24		0.2001	0.2401		0.707100	1.00
40	7	2	0.20	0.24		0.2000	0.2400		0.993730	1.00
1	9	1	0.20			0.2000			0.707356	1.00
40	9	1	0.20			0.2000			0.993730	1.00
1	9	2	0.20	0.24		0.2000	0.2400		0.706865	1.00
40	9	2	0.20	0.24		0.2000	0.2400		0.993730	1.00
1	9	3	0.20	0.24	0.28	0.2001	0.2404	0.2801	0.707106	1.00
40	9	3	0.20	0.24	0.28	0.2001	0.2403	0.2801	0.993730	1.00

True value of β: 0.707106 for $\mu = 1$ and 0.993730 for $\mu = 40$.

Thus, $d(n) \approx d = \beta^{*^2} - 1$ for $n > K$. Hence, we get from (3.13)

$$\mathbf{u}(n+1) \approx \mathbf{u}(n) - 2h\mathbf{B}\mathbf{w}(n) \qquad \forall n > K \qquad (3.14)$$

where $\mathbf{B} = \mathbf{R} + 2\mu d\mathbf{I}_N$.

Now, writing the Taylor's series expansion for $f(u)$ (cf. (2.2)) evaluated at $u = 0$, and truncating the series at fourth order term, we obtain

$$w = f(u) \approx \frac{u}{2} - \frac{u^3}{24}. \qquad (3.15)$$

Substituting (3.15) in (3.14), we get

$$\mathbf{u}(n+1) \approx \mathbf{C}\mathbf{u}(n) + \frac{h}{12}\mathbf{B}\tilde{\mathbf{u}}(n) \qquad (3.16)$$

where $\mathbf{C} = \mathbf{I}_N - h\mathbf{B}$ and $\tilde{\mathbf{u}}(n)$ is a vector with kth element given by $\tilde{u}_k(n) = u_k^3(n)$. Iterating (3.16) from K to n and using (3.3), we obtain

$$\mathbf{u}(n) \approx \sum_{j=1}^{N} \{1 - h(\lambda_j + 2\mu d)\}^{n-K} \mathbf{e}_j \mathbf{e}_j^T \mathbf{u}(K)$$
$$+ \sum_{i=0}^{n-1-K} \sum_{j=1}^{N} \frac{h}{12} \{1 - h(\lambda_j + 2\mu d)\}^i$$
$$\cdot (\lambda_j + 2\mu d)\mathbf{e}_j \mathbf{e}_j^T \tilde{\mathbf{u}}(n - 1 - i). \qquad (3.17)$$

Examining each term in (3.17) and using Theorem 2, we conclude that for convergence h should satisfy

$$|1 - h(\lambda_j + 2\mu d)| = 1 \qquad \forall j = 2P + 1, \cdots, N \qquad (3.18)$$
$$|1 - h(\lambda_j + 2\mu d)| < 1 \qquad \forall j = 1, \cdots, 2P \qquad (3.19)$$

which on simplification result in

$$\lambda_{\min} = -2\mu d = 2\mu(1 - \beta^{*^2}) \qquad (3.20)$$
$$0 < h < \frac{2}{\lambda_j - \lambda_{\min}} \qquad \forall j = 1, \cdots, 2P. \qquad (3.21)$$

Since the bounds on h must be satisfied for all eigenvalues, λ_1 to λ_{2P}, we replace λ_j in (3.21) with the maximum eigenvalue $\lambda_{\max}(= \lambda_1)$, thus obtaining

$$0 < h < \frac{2}{\lambda_{\max} - \lambda_{\min}} \qquad (3.22)$$

Observe that this result is same as that reported in [5]. It may be easily verified that even if we include the higher order terms in (3.15), the bounds on h would remain the same as (3.22). Thus, no simplifying assumption needs to be made on the sigmoidal nonlinearity. Since the eigenvalues of $\mathbf{R}$ are not known *a priori*, we suggest the following practical bounds for h

$$0 < h < \frac{2}{\text{Trace}(\mathbf{R})} \qquad (3.23)$$

IV. SIMULATION RESULTS

For the data described by (1.1), the asymptotic autocorrelation matrix $\mathbf{R}$ is given by

$$R(i, j) = \sum_{l=1}^{P} \frac{\alpha_l^2}{2} \cos(\omega_l k) + \sigma^2 \delta_{ij}$$

where $k = |i - j|, i, j = 1, \cdots, N$. The system of differential equations was solved numerically, with the integration time-step h chosen according to (3.23). The iterations were stopped when the norm of the difference between the consecutive solution vectors was less than a predetermined threshold, δ(i.e., $\|\mathbf{w}(n+1) - \mathbf{w}(n)\|_2 < \delta$). Then, frequencies of the sinusoids were computed from the roots of the polynomial (formed using the estimated minimum eigenvector) which were closest to the unit circle. If $\hat{\mathbf{w}}$ denotes the estimated minimum eigenvector, then the minimum eigenvalue was estimated as

$$\hat{\lambda}_{\min} = \frac{\hat{\mathbf{w}}^T \mathbf{R} \hat{\mathbf{w}}}{\hat{\mathbf{w}}^T \hat{\mathbf{w}}}$$

and $\|\hat{\mathbf{w}}\|_2$ was taken as the estimate of β.

In the simulations, we chose $\sigma^2 = 1$ (giving $\lambda_{\min} = 1$) and $\delta = 10^{-6}$. For a fixed h, the estimated values of the frequencies of the sinusoids, $\hat{\lambda}_{\min}$ and $\hat{\beta}$, for different values of N, P and μ are given in the table. We note the following from the results.

When μ is large, $\hat{\beta}$ is closer to unity as predicted by the cost function (2.1) and Theorem 2. The estimated value of $\lambda_{\min}$ is same as the true value and the norm of the solution vector, $\hat{\beta}$, is very close to the theoretical value given by (3.20).

We also verified the existence of multiple solutions for the case $N > 2P + 1$ by using different initial conditions ($\mathbf{w}(0)$). For $\mu \leq (\lambda_{\min}/2)$, behavior of the system was erroneous.

V. CONCLUSIONS

The problem of estimating the frequencies of real sinusoids corrupted with white noise using the Pisarenko's harmonic retrieval method has been recast into the neural network framework. An analysis investigating the nature of the minimizers of the energy function and the convergence of the numerical technique, used for solving the network dynamics, is presented. Results of the analysis are supported by simulations. Though we considered the symptomatic case in the paper, the approach can be easily extended to the finite data case.

## ACKNOWLEDGMENT

REFERENCES

[1] J. J. Hopfield and D. W. Tank, "Neural computations of decisions in optimization problems," *Biological Cybern.*, vol. 52, pp. 141–152, 1985.

[2] P. A. Thompson, "An adaptive spectral analysis technique for unbiased frequency estimation in the presence of white noise," in *Proc. 13th Asilomar Conf. Circuits, Syst., Comput.* (Pacific Grove, CA), Nov. 1979, pp. 529–533.

[3] V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. Royal Astron. Soc.*, pp. 347–366, 1973.

[4] V. U. Reddy, B. Egardt, and T. Kailath, "Least squares type algorithm for adaptive implementation of Pisarenko's harmonic retrieval method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 399–405, June 1982.

[5] M. G. Larimore, "Adaption convergence of spectral estimation based on Pisarenko harmonic retrieval," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 955–962, Aug. 1983.

[6] D. R. Fuhrmann and B. Liu, "Rotational search methods for adaptive Pisarenko harmonic retrieval," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1550–1565, Dec. 1986.

[7] D. G. Luenberger, *Linear and Non-linear Programming, .* Reading, MA: Addison-Wesley, 1978, pp. 366–369.

[8] E. Oja, "A Simplified neuron model as a principal component analyzer," *J. Math. Biology*, vol. 15, pp. 267–273, 1982.

[9] P. Baldi and K. Hornik, "Neural networks and principal component analysis: Learning from examples without local minima," *Neural Networks*, vol. 2, pp. 53–58, 1989.

[10] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward neural network," *Neural Networks*, vol. 2, pp. 459–473, 1989.

[11] S. Y. Kung and K. I. Diamantaras, "A neural network learning algorithm for adaptive principal component extraction (APEX)," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 861–864, 1990.

[12] S. Y. Kung "Constrained principal component analysis via an orthogonal learning network," in *Proc. Int. Symp. Circuits Syst.*, 1990, pp. 719–722.

[13] Y. Chauvin, "Principal component analysis by gradient descent on a constrained linear Hebbian cell," in *Proc. Int. Joint Conf. Neural Networks*, 1989, pp. 1.373–1.380.
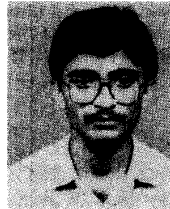
George Mathew received the B.E. degree in electronics and communication engineering from Karnataka Regional Engineering College, Surathkal, India, in 1987, and the M.Sc.(Engg.) degree in electrical communication engineering from the Indian Institute of Science (IISc), Bangalore, in 1989. Currently, he is working toward the Ph.D. degree in electrical communication engineering at IISc, Bangalore.

His research interests are adaptive algorithms and neural networks.
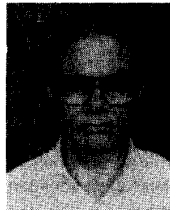
V. U. Reddy received the B.E. and M.Tech. degrees in electronics and communication engineering from Osmania University and the Indian Institute of Technology (IIT), Kharagpur, in 1962 and 1963, respectively, and Ph. D. degree in electrical engineering from the University of Missouri, Columbia, in 1971.

He was an Assistant Professor at IIT, Madras, India, during 1972–1976 and Professor at IIT, Kharagpur, during 1976–1979. During 1979–1982 and 1986–1987, he was a Visiting Professor at the Department of Electrical Engineering, Stanford University, Stanford, CA. In April 1982, he joined Osmania University as a Professor and was the Founder-Director of the Research and Training Unit for Navigational Electronics, funded by the Department of Electronics, Government of India. Since April 1988, he has been with the Indian Institute of Science, Bangalore, as a Professor of Electrical Communication Engineering and is presently its Chairman. He has served as a consultant in signal processing to Avionics Design Bureau of Hindustan Aeronautics Limited, Hyderabad, and to Central Research Laboratory, Bharat Electronics Limited, Bangalore. His recent research interests are in sensitivity analysis of high-resolution algorithms, adaptive algorithms, adaptive arrays and wavelet transforms.

Dr. Reddy is a Fellow of the Indian Academy of Sciences, Indian National Academy of Engineering and Indian National Science Academy, and Fellow of the Institution of Electronics and Telecommunications Engineers (IETE), India. He received the S.K. Mitra Memorial Award (1989) from IETE for the best research paper.