Joint Time Delay and Frequency Estimation via State-Space Realization

Y. Wu, H. C. So, *Member, IEEE*, and P. C. Ching, *Senior Member, IEEE*

Abstract—By applying a two-dimensional parameter estimation method proposed by Viberg and Stoica, we develop a subspace method for estimating the differential delay of a sinusoidal signal received at two separated sensors as well as the sinusoidal frequencies. Using state-space realization, the time delay and frequency estimates are obtained from the state transition and observation matrices. Performance evaluation via computer simulations is included to demonstrate the effectiveness of the proposed algorithm.

Index Terms—Delay estimation, frequency estimation, state-space, subspace algorithm.

I. INTRODUCTION

TIME DELAY estimation between signals received at spatially separated sensors [1], [2] and sinusoidal frequency estimation [3], [4] are two frequently addressed problems in the signal processing literature. Recently, the problem of joint time delay and frequency estimation of sinusoidal signals has also attracted considerable attention. Application examples for this work include speech enhancement and pitch estimation using a microphone array [5], [6], synchronization in code division multiple access (CDMA) systems [7], analysis of thalamocortical seizure pathways [8] and fequency-shift keying (FSK) demodulation using multiple segments [9].

Let the discrete-time sinusoidal signals received at two sensors be

$$r_1(n) = s(n) + q_1(n)$$
$$r_2(n) = s(n - D) + q_2(n), \quad n = 0, 1, \ldots, N - 1 \quad (1)$$

where

$$s(n) = \sum_{m=1}^{P} \alpha_m \exp(j\omega_m n). \quad (2)$$

The source signal $s(n)$ is modeled by a sum of P complex sinusoids where the amplitudes $\{\alpha_m\}$ are unknown complex-valued

Y. Wu is with Laboratory for Radar Signal Processing, Xidian University, Xian 710071, China.

H. C. So is with the Department of Computer Engineering and Information Technology, City University of Hong Kong, Hong Kong (e-mail: ithcso@cityu.edu.hk).

P. C. Ching is with Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (e-mail: pcching@ee.cuhk.edu.hk).

constants, and the normalized radian frequencies $\{\omega_m\}$ are distinct. It is assumed that P is known *a priori* or an accurate estimate of P has been obtained [10]. The additive noises $q_1(n)$ and $q_2(n)$ are uncorrelated zero-mean complex white Gaussian processes with variances σ_q^2. The parameter D represents the difference in arrival times at the two receivers and N is the number of samples collected at each channel. Our goal is to estimate both the time difference of the received signals and the frequencies of their constituent components.

When $s(n)$ is a single sinusoid, a discrete-time Fourier transform based method [11] has been devised for optimum time delay and frequency estimation. For $P > 1$, Sherman *et al.* [8] has presented an ESPRIT algorithm [12] to estimate D while a generalized Yule–Walker solution is suggested in [3] to determine $\{\omega_m\}$ separately. Recently, a subspace-based method [6] has been developed for joint delay and frequency estimation and the estimates are obtained using the eigenvalues and eigenvectors of a matrix derived from the covariance matrices of the received signals. Unlike standard subspace approaches where the frequencies are estimated from the eignenvalues, this technique calculates the frequencies from the eigenvectors, and as a result its estimation performance is limited. In this paper, we first formulate the problem using a state-space realization [13] and then utilize the subspace technique to achieve joint parameter estimation. The state-space model for multiple sinusoidal frequency estimation was first introduced in [14], which was motivated by the fact that state-space parameterization enables reduction of parameter sensitivity [15]. This model has been extended to direction-of-arrival (DOA) estimation [16] as well as joint DOA and frequency estimation [17] using array data, where the second method is referred to as the Viberg–Stoica method. Our approach in obtaining the joint delay and frequency estimation is based on the Viberg–Stoica method and will be explained in detail in Section II. In short, the frequency estimates are given by the eigenvalue phases of a state transition matrix while the delay is determined using the diagonal elements of an observation matrix as well as the estimated frequencies. In Section III, simulation results are included to evaluate the estimator performance for different parameter settings and comparisons with the Cramér–Rao lower bound (CRLB) are also made. Finally, conclusions are drawn in Section IV.

II. DEVELOPMENT OF PROPOSED METHOD

The development of the proposed method consists of two steps, namely, formulating the joint time delay and frequency estimation problem using state-space realization, and then applying the Viberg–Stoica method for parameter estimation.

A. State-Space Model Formulation

Using the received signals, we form the following set of vectors:

$$\mathbf{X_1}(k) = [r_1(k), r_1(k+1), \ldots, r_1(k+M-1)]^T$$
$$\mathbf{X_2}(k) = [r_2(k), r_2(k+1), \ldots, r_2(k+M-1)]^T \quad (3)$$

where $k = 0, 1, \ldots, K-1$ with $K = N - M + 1$; M is the length of each vector; and T denotes the transpose operation. Note that both M and K should be larger than P.

Substituting (1) and (2) into (3) gives

$$\mathbf{X_1}(k) = \mathbf{A}(\omega)\mathbf{S}(k) + \mathbf{Q}_1(k)$$
$$\mathbf{X_2}(k) = \mathbf{A}(\omega)\mathbf{\Delta}(\omega, D)\mathbf{S}(k) + \mathbf{Q}_2(k) \quad (4)$$

where

$$\mathbf{A}(\omega) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \cdots & e^{j\omega_P} \\ \vdots & \vdots & \vdots & \vdots \\ e^{j\omega_1(M-1)} & e^{j\omega_2(M-1)} & \cdots & e^{j\omega_P(M-1)} \end{bmatrix}$$

$$\mathbf{S}(k) = [\alpha_1 e^{j\omega_1 k}, \alpha_2 e^{j\omega_2 k}, \ldots, \alpha_P e^{j\omega_P k}]^T$$
$$\mathbf{Q}_1(k) = [q_1(k), q_1(k+1), \ldots, q_1(k+M-1)]^T$$
$$\mathbf{Q}_2(k) = [q_2(k), q_2(k+1), \ldots, q_2(k+M-1)]^T$$
$$\mathbf{\Delta}(\omega, D) = \mathrm{diag}(e^{-jD\omega_1}, e^{-jD\omega_2}, \ldots, e^{-jD\omega_P}). \quad (5)$$

Using $\mathbf{S}(k)$ as the state vector, and grouping $\mathbf{X_1}(k)$ and $\mathbf{X_1}(k)$ into one vector, namely, $\mathbf{X}(k) = [\mathbf{X_1}^T(k)\mathbf{X_2}^T(k)]^T$, we can formulate a state-space model for the array output as

$$\mathbf{S}(k+1) = \mathbf{\Phi}\mathbf{S}(k) \quad \text{and} \quad \mathbf{X}(k) = \mathbf{B}\mathbf{S}(k) + \mathbf{E}(k) \quad (6)$$

where

$$\mathbf{\Phi} = \mathrm{diag}(e^{j\omega_1}, e^{j\omega_2}, \ldots, e^{j\omega_P})$$
$$\mathbf{B} = \begin{bmatrix} \mathbf{A}(\omega) \\ \mathbf{A}(\omega)\mathbf{\Delta}(\omega, D) \end{bmatrix}$$
$$\mathbf{E}(k) = \begin{bmatrix} \mathbf{Q}_1(k) \\ \mathbf{Q}_2(k) \end{bmatrix}. \quad (7)$$

In the state-space terminology, $\mathbf{\Phi}$ and $\mathbf{B}$ are referred to as the state transition matrix and observation matrix, respectively. Our estimation procedure first finds $\mathbf{\Phi}$ from which the frequency estimates can be obtained. Estimation of $\mathbf{B}$ will then be followed and finally, D is acquired with the use of the estimated frequencies.

B. Estimation via Applying Viberg–Stoica Method

The estimation algorithm is based on the canonical variables [13], which employs the structure of the cross correlation between future and past data. Following the framework of the Viberg–Stoica method [17], we construct a future output vector

$$\mathbf{X}_L(k) = [\mathbf{X}^T(k), \mathbf{X}^T(k+1), \ldots, \mathbf{X}^T(k+L-1)]^T,$$
$$k = 1, 2, \ldots, K - L + 1 \quad (8)$$

which is a $2ML$-vector with $L \geq 2$. Using (6), $\mathbf{X}_L(k)$ is expressed as

$$\mathbf{X}_L(k) = \mathbf{\Omega}_L\mathbf{S}(k) + \mathbf{E}_L(k) \quad (9)$$

where

$$\mathbf{\Omega}_L = [\mathbf{B}^T \quad (\mathbf{B}\mathbf{\Phi})^T \cdots (\mathbf{B}\mathbf{\Phi}^{L-1})^T]^T \quad (10)$$

and

$$\mathbf{E}_L(k) = [\mathbf{E}^T(k), \mathbf{E}^T(k+1), \ldots, \mathbf{E}^T(k+L-1)]^T. \quad (11)$$

It is required that $\mathbf{\Omega}_{L-1}$ must have full-column rank of P, and a necessary condition for it is $2M(L-1) \geq P$. If $P \leq 2M$, a sufficient condition is that $\mathbf{B}$ has full-column rank. The cross-covariance matrix between $\mathbf{X}_L(k)$ and the past output vector $\mathbf{X}_L(k-L)$ is

$$\mathbf{R}_{X_L}(L) \triangleq E\{\mathbf{X}_L(k)\mathbf{X}_L^H(k-L)\} = \mathbf{\Omega}_L\mathbf{R}_S\mathbf{\Omega}_L^H \quad (12)$$

where $\mathbf{R}_S = \mathrm{diag}(|\alpha_1|^2 e^{j\omega_1 L}, |\alpha_2|^2 e^{j\omega_2 L}, \ldots, |\alpha_P|^2 e^{j\omega_P L})$, because the noise components are uncorrelated. Hence, $\mathbf{R}_{X_L}(L)$ has low rank equal to P, and its column space coincides with that of $\mathbf{\Omega}_L$. A set of orthonormal basis vectors for the column space can be obtained from the P principal left singular vectors of $\mathbf{R}_{X_L}(L)$. In practice, the vectors are computed from the singular value decomposition (SVD) of the sample cross covariance

$$\hat{\mathbf{R}}_{X_L}(L) = \frac{1}{K} \sum_{k=L}^{K-L+1} \mathbf{X}_L(k)\mathbf{X}_L^H(k-L) = \sum_{l=1}^{2ML} \gamma_l \mathbf{u}_l\mathbf{v}_l^H \quad (13)$$

where $\{\gamma_l\}$ represent the singular values arranged in nonincreasing order, while $\mathbf{u}_l$ and $\mathbf{v}_l$ denote the left and right singular vectors, respectively. When N and/or the signal-to-noise ratio (SNR) is sufficiently large, the range of $\hat{\mathbf{\Omega}} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_P]$ can be well approximated by the range of $\mathbf{\Omega}_L$, which implies

$$\hat{\mathbf{\Omega}} \approx \mathbf{\Omega}_L\mathbf{T} \quad (14)$$

where $\mathbf{T}$ is an unknown full-rank $P \times P$ state transformation matrix. In fact, $\hat{\mathbf{\Omega}}$ is consistent in the sense that $\lim_{N\to\infty} \hat{\mathbf{\Omega}} = \mathbf{\Omega}_L\mathbf{T}$. This means that from $\hat{\mathbf{\Omega}}$, we can find the corresponding transformed state matrices for $\mathbf{B}$ and $\mathbf{\Phi}$

$$\mathbf{B}_T = \mathbf{B}\mathbf{T} \quad \mathbf{\Phi}_T = \mathbf{T}^{-1}\mathbf{\Phi}\mathbf{T} \quad (15)$$
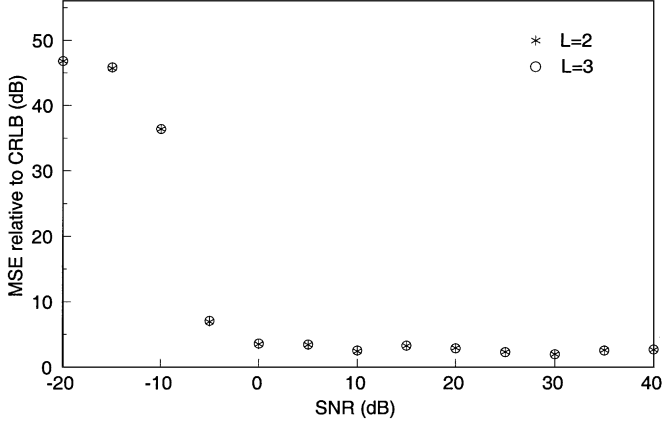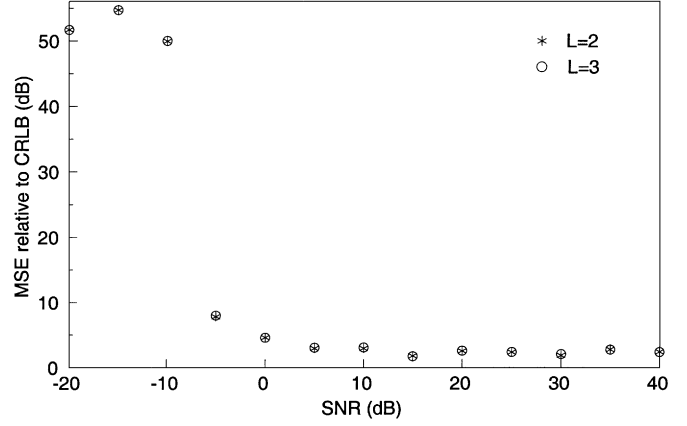
and they are estimated as [18]

$$\hat{\mathbf{B}}_T = \hat{\mathbf{\Omega}}_{1:1} \quad (16)$$

and

$$\hat{\mathbf{\Phi}}_T = \hat{\mathbf{\Omega}}_{1:L-1}^{\#}\hat{\mathbf{\Omega}}_{2:L} \quad (17)$$

where $\hat{\mathbf{\Omega}}_{k:l}$ denotes the block rows from k through l, and $\#$ is the pseudoinverse operation.

Fig. 1.   MSE relative to CRLB of $\hat{\omega}_1$ versus SNR.

Fig. 2.   MSE relative to CRLB of $\hat{\omega}_2$ versus SNR.

Since $\mathbf{\Phi}$ is a diagonal matrix with eigenvalues $e^{j\omega_m}$, $m = 1, 2, \ldots, P$, estimation of $\mathbf{\Phi}$ is achieved via diagonalizing $\hat{\mathbf{\Phi}}_T$ as

$$\hat{\mathbf{\Phi}}_T = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} \tag{18}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_P)$. From (15) and (18), estimates of $\mathbf{\Phi}$ and $\mathbf{T}$ are given by $\mathbf{\Lambda}$ and $\mathbf{U}^{-1}$, respectively. As a result, consistent frequency estimates are computed as

$$\hat{\omega}_m = \angle\lambda_m, \quad m = 1, 2, \ldots, P. \tag{19}$$

On the other hand, $\mathbf{B}$ is estimated as $\hat{\mathbf{B}} = \hat{\mathbf{B}}_T\mathbf{U}$. Since the upper and lower parts of $\mathbf{B}$ are $\mathbf{A}(\omega)$ and $\mathbf{A}(\omega)\mathbf{\Delta}(\omega, D)$, respectively, we have $\hat{\mathbf{B}}_{1:M}\hat{\mathbf{\Delta}}(\omega, D) \approx \hat{\mathbf{B}}_{M+1:2M}$ where $\hat{\mathbf{\Delta}}(\omega, D)$ represents the estimate of $\mathbf{\Delta}(\omega, D)$. A least squares solution for $\hat{\mathbf{\Delta}}(\omega, D)$ is $\hat{\mathbf{B}}_{1:M}^{\#}\hat{\mathbf{B}}_{M+1:2M}$, and finally the time delay can be estimated using all diagonal elements of $\hat{\mathbf{\Delta}}(\omega, D)$ and $\{\hat{\omega}_m\}$

$$\hat{D} = \frac{\sum_{m=1}^{P} \angle\hat{\mathbf{\Delta}}_{m,m}(\omega, D)}{-\sum_{m=1}^{P} \hat{\omega}_m}. \tag{20}$$

Notice that when N is much larger than M and L, the computational complexity of the algorithm is in the order of M^2L^2N.

III. SIMULATION RESULTS

Computer simulations were conducted to evaluate the joint time delay and frequency estimation performance of the proposed method in the presence of white Gaussian noise. The source signal $s(n)$ was a sinusoidal signal of the form $s(n) = \alpha_1 e^{j\omega_1 n} + \alpha_2 e^{j\omega_2 n}$ with $\alpha_1 = \alpha_2 = 1/\sqrt{2}$, $\omega_1 = 0.2\pi$ rad/s and $\omega_2 = 0.4\pi$ rad/s. The sampling interval was 1s and the time delay D was selected to be 1.7s. Different SNRs were obtained by proper scaling of the noise sequences. The number of samples was $N = 200$ and the value of M was chosen to be 25. All results provided were averages of 200 independent runs.

Figs. 1 and 2 plot the mean square error (MSE) of the frequency estimates relative to the CRLB, which was defined as MSE/CRLB, versus SNR for $L = 2$ and $L = 3$. It can be seen that the proposed algorithm performed similarly for both Ls and the variances of $\hat{\omega}_1$ and $\hat{\omega}_2$ were only larger than the corresponding CRLBs by around 2 dB when SNR > -5 dB.
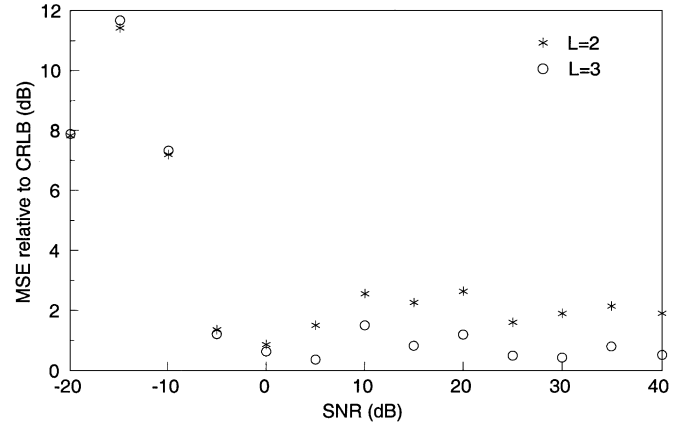
Fig. 3.   MSE relative to CRLB of D versus SNR.

While the results of MSE relative to the CRLB for delay estimation are shown in Fig. 3. We observe that the algorithm with $L = 3$ was superior to that of $L = 2$ and degraded from the CRLB, both by approximately 1 dB at SNR ≥ 5 dB, although the two settings could give accurate estimation performance for SNR ≥ -5 dB. On the other hand, the computational load for $L = 3$ was around 9/4 times of that of $L = 2$. Hence, we see that there is a trade-off between computational complexity and delay estimation accuracy. It is noteworthy that the algorithm with $L = 2$ is closely related to [12], and thus it can be considered as a ESPRIT-type algorithm. Alternatively, our approach can be viewed as a generalized algorithm.

IV. CONCLUDING REMARKS

A subspace algorithm based on state-space realization has been proposed for joint time delay and frequency estimation of sinusoidal signals received at two separated sensors. The frequency estimates are obtained directly from the eigenvalues of the state transition matrix while the delay is determined using the observation matrix and the estimated frequencies. Performance evaluation via computer simulations shows that the frequency and delay variances of the proposed approach can be close to the corresponding performance bounds for sufficiently high SNR. We have also applied the algorithm for joint time delay and pitch estimation using voiced speech at a microphone array, and the findings are reported in [19].

REFERENCES

[1] "Special issue on time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, pp. 582–587, June 1981.

[2] G. C. Carter, *Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test and Evaluation Engineers*. Piscataway, NJ: IEEE Press, 1993.

[3] S. M. Kay, *Modern Spectral Estimation: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[4] P. Stoica, "List of references on spectral line analysis," *Signal Process.*, vol. 31, no. 3, pp. 329–340, Apr. 1993.

[5] X. Qian and R. Kumaresan, "Joint estimation of time delay and pitch of voiced speech signals," in *Conf. Rec. 29th Asilomar Conf. Signals, Systems Computers*, vol. 1, Pacific Grove, CA, Nov. 1995, pp. 735–739.

[6] G. Liao, H. C. So, and P. C. Ching, "Joint time delay and frequency estimation of multiple sinusoids," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing*, Salt Lake, UT, May 2001, pp. 3121–3124.

[7] S. R. Dooley and A. K. Nandi, "Adaptive time delay and frequency estimation for digital signal synchronization in CDMA systems," in *Conf. Rec. 32th Asilomar Conf. Signals, Systems & Computers*, vol. 2, Pacific Grove, CA, Nov. 1998, pp. 1838–1842.

[8] D. L. Sherman, Y. C. Tsai, L. A. Rossell, M. A. Mirski, and N. V. Thakor, "Narrowband delay estimation for thalamocortical epileptic seizure pathways," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing*, vol. 5, Detroit, MI, May 1995, pp. 2939–2942.

[9] J. A. Sills and Q. R. Black, "Frequency estimation from short pulses of sinusoidal signals," *Proc. IEEE MILCOM '96*, vol. 3, pp. 979–983, 1996.

[10] H. T. Wu, J. F. Yang, and F. K. Chen, "Source number estimators using transformed gerschgorin radii," *IEEE Trans. Signal Processing*, vol. 43, pp. 1325–1333, June 1995.

[11] H. C. So, "Delay estimation for sinusoidal signals," *Proc. Inst. Elect. Eng.—Radar, Sonar, Navigat.*, vol. 148, no. 6, pp. 318–324, Dec. 2001.

[12] R. Roy and T. Kailath, "ESPRIT—Estimation of Signal Parameter via Rotational Invariance Techniques," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, pp. 984–995, July 1989.

[13] M. Aoki, *State Space Modeling of Time Series*. Berlin, Germany: Springer-Verlag, 1990.

[14] S. Y. Kung, K. S. Arun, and B. D. Rao, "State space and SVD based approximation methods for the harmonic retrieval problem," *J. Opt. Soc. Amer.*, vol. 73, pp. 1799–1811, Dec. 1983.

[15] B. D. Rao, "Sensitivity considerations in state-space model-based harmonic retrieval methods," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, pp. 1789–1794, Nov 1989.

[16] S. Prasad and B. Chandna, "Direction of arrival estimation using stochastic model order reduction via state space modeling," *Signal Process.*, vol. 23, no. 2, pp. 157–177, May 1991.

[17] M. Viberg and P. Stoica, "A computationally efficient method for joint direction finding and frequency estimation in colored noise," in *Conf. Rec. 32nd Asilomar Conf. Signals, Systems Computers*, vol. 2, Pacific Grove, CA, Nov 1998, pp. 735–739.

[18] S. Y. Kung, "A new identification and model reduction algorithm via singular value decomposition," in *Proc. 12th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov 1978, pp. 705–714.

[19] L. Y. Ngan, Y. Wu, H. C. So, and P. C. Ching, "Joint time delay and pitch estimation for speaker localization," in *Proc. ISCAS*, Bangkok, Thailand, 2003, pp. 722–725.