way that each position provides a time-adjusted null constraint to the other in a linearly constrained beamformer such as the GSC [1]. This multisource structure [13] performs better than MUSIC in both mobile and immobile cases, as shown in Fig. 3(b). It is one of the issues reported in [13], among other generalizations [14].

REFERENCES

[1] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, no. 1, pp. 27–34, Jan. 1982.

[2] B. Widrow, K. M. Duvall, R. P. Gooch, and W. C. Newman, "Signal cancellation phenomena in adaptive antennas: Causes and cures," *IEEE Trans. Antennas Propagat.*, vol. AP-30, no. 3, pp. 469–478, May 1982.

[3] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.

[4] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 5, pp. 1235–1247, Oct. 1983.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 3, pp. 276–280, Mar. 1986.

[6] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. ICASSP'83*, Boston, vol. 1, Apr. 14–16, 1983, pp. 336–339.

[7] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, July 1989.

[8] B. Yang, "Subspace tracking based on the projection approach and the recursive least squares method," in *Proc. ICASSP'93*, Minneapolis, MN, vol. IV, Apr. 27–30, 1993, pp. 145–148.

[9] A. Eriksson, P. Stoica, and T. Söderström, "On-line subspace algorithms for tracking moving sources," *IEEE Trans. Signal Processing*, vol. 42, no. 9, pp. 2319–2330, Sept. 1994.

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

[11] S. Affes, S. Gazor, and Y. Grenier, "Robust adaptive beamforming via LMS-like target tracking," in *Proc. ICASSP'94*, Adelaide, Australia, vol. IV, Apr. 19–22, 1994, pp. 269–273.

[12] ——, "Analysis of LMS-like source tracking for robust adaptive beamforming," in *Proc. EURASIP EUSIPCO'94*, Edinburgh, Scotland, vol. II, Sept. 13–16, 1994, pp. 760–763.

[13] ——, "An algorithm for multisource beamforming and multitarget tracking," *IEEE Trans. Signal Processing*, this issue, pp. 1512–1522.

[14] S. Gazor, S. Affes, and Y. Grenier, "Wideband multi-source beamforming with adaptive array location calibration and direction finding," in *Proc. ICASSP'95*, Detroit, MI, vol. III, May 8–12, 1995, pp. 1904–1907.

Connections Between the Least-Squares and the Subspace Approaches to Blind Channel Estimation

Hanks H. Zeng and Lang Tong

***Abstract*—In this correspondence, we study the connections between the least-squares and the subspace approaches to blind channel estimation. By examining the properties and connections of the so-called multichannel filtering and data selection transforms, we establish a relationship between the identification equations used in the two approaches. Next, it is shown that the least-squares and subspace estimators are identical for the case when there are two subchannels. In general, the two algorithms are different in their utilization of the noise subspace.**

I. INTRODUCTION

The so-called blind channel identification, i.e., identifying a channel using only the channel output, has attracted increasing research attention in recent years. Since the publication of [7], several interesting eigenstructure-based approaches [2], [4], [5] have been proposed. In some simulations, there is a significant performance improvement over the algorithm proposed in [7] and [8]. Much of the performance gain can be credited to the exploitation of the special structure of the so-called filtering transform. Under the identifiability condition [6]–[8], the first two such algorithms are the *subspace approach* (SS) proposed by Moulines *et al.* [3], [5] and the *least-squares approach* (LS) proposed in [2]. Slock also derived similar results using linear prediction techniques [5]. The SS method is based on the following two key results: i) The signal subspace uniquely determines the channel impulse response, and ii) the channel vector is orthogonal to the filtering transform of the noise subspace. The LS approach, on the other hand, is derived by the exploiting the single-input multiple-output nature of the identification procedure.

In this correspondence, we investigate connections between the SS and LS approaches. Assuming that the same data window is used, the LS and SS approaches can be derived from the same covariance matrix of the channel outputs. The identification equations and the uniqueness of their solutions are shown in the same framework. The approach presented here offers a unified view of the two methods. The second result of this paper is to show that the two estimators are in fact *identical* for the special case involving two subchannels. This case is of particular importance since it corresponds to the $T/2$-fractionally sampled channel that is popular in communication applications. Differences between the two approaches are also revealed in our new derivation.

II. THE MODEL

A. Channel Model and Assumptions

We consider the blind channel identification of a discrete-time single-input multiple-output model given by

$$x_i(t) \triangleq \sum_{n=0}^{L} h_i(n) w_{t-n},$$

Manuscript received April 7, 1995; revised December 1, 1995. This work was supported, in part, by the National Science Foundation under Contract NCR-9321813 and by the Advanced Research Projects Agency monitored by the Federal Bureau of Investigation under Contract no. J-FBI-94-221. The associate editor coordinating the review of this paper and approving it for publication was Dr. Zhi Ding.

The authors are with the Department of Electrical and Systems Engineering, The University of Connecticut, Storrs, CT 06269-3157 USA.

Publisher Item Identifier S 1053-587X(96)03960-8.

TABLE I
MODEL DEFINITION AND AN EXAMPLE, SUPERSCRIPT H AND T DENOTE CONJUGATE TRANSPOSE AND TRANSPOSE, RESPECTIVELY. $(\mathbf{h}_i)$ IS A $(L+1)\times(2L+1)$ MATRIX. $\mathcal{H}_{M,L}(\mathbf{h})$ IS A $M(L+1)\times(2L+1)$ MATRIX

Notation	Definition	Example
M	the number of subchannels.	2
$h_i(z)$	$h_i(z)=h_i(0)+h_i(1)z^{-1}+\cdots+h_i(L_i)z^{-L_i}$	$h_1(z)=0.4+0.6z^{-1},\quad h_2(z)=0.5+0.3z^{-1}$
L	$L=\max_i\{deg\{h_i(z)\}\}$	1
$\mathbf{h}_i$	$\mathbf{h}_1=[h_i(0),\cdots,h_i(L)]^T\in C^{L+1}$	$\mathbf{h}_1=[0.4,0.6]^T,\quad \mathbf{h}_2=[0.5,0.3]^T$
$\mathbf{h}$	$\mathbf{h}=[\mathbf{h}_1^H,\cdots,\mathbf{h}_M^H]^H$	$\mathbf{h}=[0.4,0.6\ \vdots\ 0.5,0.3]^T$
$\bar{\mathbf{x}}_i(t)$	$\bar{\mathbf{x}}_i(t)=[x_i(t),\cdots,x_i(t-L)]^T$	$\bar{\mathbf{x}}_i(t)=[x_i(t),x_i(t-1)]^T,\ i=1,2$
$\mathbf{x}(t)$	$\mathbf{x}(t)=[\bar{\mathbf{x}}_1^H,\cdots,\bar{\mathbf{x}}_M^H]^H$	$\mathbf{x}(t)=[x_1(t),x_1(t-1)\ \vdots\ x_2(t),x_2(t-1)]^T$
$\bar{\mathbf{y}}_i(t)$	$\bar{\mathbf{y}}_i(t)=[y_i(t),\cdots,y_i(t-L)]^T$	$\bar{\mathbf{y}}_i(t)=[y_i(t),y_i(t-1)]^T,\quad i=1,2$
$\mathbf{y}(t)$	$\mathbf{y}(t)=[\bar{\mathbf{y}}_1^H,\cdots,\bar{\mathbf{y}}_M^H]^H$	$\mathbf{y}(t)=[y_1(t),y_1(t-1)\ \vdots\ y_2(t),y_2(t-1)]^T$
$\mathbf{w}(t)$	$\mathbf{w}(t)=[w_t,\cdots,w_{t-2L}]^T$	$\mathbf{w}(t)=[w_t,w_{t-1},w_{t-2}]^T$
$\mathcal{F}_L(\mathbf{h}_i)$	$\mathcal{F}_L(\mathbf{h}_i)=\begin{pmatrix} h_0^{(i)} & \cdots & h_L^{(i)} & & \\ & \ddots & & \ddots & \\ & & h_0^{(i)} & \cdots & h_L^{(i)}\end{pmatrix}$	$\mathcal{F}_L(\mathbf{h}_1)=\begin{pmatrix}0.4 & 0.6 & 0\\ 0 & 0.4 & 0.6\end{pmatrix}$
$\mathcal{H}_{M,L}(\mathbf{h})$	$\mathcal{H}_{M,L}(\mathbf{h})=[\mathcal{F}^H(\mathbf{h}_1)\cdots\mathcal{F}^H(\mathbf{h}_M)]^H$	$\mathcal{H}_{M,L}(\mathbf{h})=\begin{pmatrix}0.4 & 0.6 & 0\\ 0 & 0.4 & 0.6\\ 0.5 & 0.3 & 0\\ 0 & 0.5 & 0.3\end{pmatrix}$

$$y_i(t) = x_i(t)+n_i(t) \tag{1}$$

where the $\{h_i(t)\}$ are impulse responses of subchannels to be estimated using observations $\{y_i(t)\}$. Under the notation given in Table I, a vector representation of (1) is given by

$$\begin{aligned}\mathbf{x}(t) &= \mathcal{H}_{M,L}(\mathbf{h})\mathbf{w}(t),\\ \mathbf{y}(t) &= \mathbf{x}(t)+\mathbf{n}(t).\end{aligned} \tag{2}$$

We impose the following assumptions in the sequel:

A1) Subchannels do not share common zeros, i.e., $\bigcap_{i=1}^{M}\mathcal{Z}[h_i(z)]=\emptyset$, where $\mathcal{Z}[h_i(z)]$ denotes the zeros of the polynomial $h_i(z)$.

A2) The input sequence $\{w_t\}$ is wide-sense stationary with zero mean and correlation $\mathbf{R}_w \triangleq E\{\mathbf{w}(t)\mathbf{w}(t)^H\}>0$.

A3) The noise $n(t)$ is white with zero mean and variance σ^2.

Note that A1 ensures channel identifiability, and it also implies that $\mathcal{H}_{M,L}(\mathbf{h})$ has a full column rank [6].

B. The Multichannel Filtering and Data Selection Transforms

The multichannel filtering transform (MFT) $\mathcal{H}_{M,L}(\mathbf{h})$ plays the key role in the development of the subspace approach to blind channel identification and estimation. For the least-squares approach, on the other hand, it is the so-called data selection transform (DST) defined below that combines output from different channels to form identification equations. These two transforms have some similar properties, such as linearity and symmetry. More important, they are related by properties of orthogonality and commutativity, which form the basis of connecting the SS and LS approaches.

Let $\mathbf{B}\in C^{M(L+1)\times l}$ be a matrix consisting of vectors $\mathbf{b}_i^{(j)}=[b_i^{(j)}(0),\cdots,b_i^{(j)}(L)]^T\in C^{L+1}$

$$\mathbf{B}=\begin{pmatrix}\mathbf{B}_1\\ \vdots\\ \mathbf{B}_M\end{pmatrix}=\begin{pmatrix}\mathbf{b}_1^{(1)} & \cdots & \mathbf{b}_1^{(l)}\\ \vdots & \cdots & \vdots\\ \mathbf{b}_M^{(1)} & \cdots & \mathbf{b}_M^{(l)}\end{pmatrix}. \tag{3}$$

1) Multichannel Filtering Transform (MFT) [3]:

$$\mathcal{H}_{M,L}(\mathbf{B})\triangleq\begin{pmatrix}\mathcal{F}_L(\mathbf{b}_1^{(1)}) & \cdots & \mathcal{F}_L(\mathbf{b}_1^{(l)})\\ \vdots & \cdots & \vdots\\ \mathcal{F}_L(\mathbf{b}_M^{(1)}) & \cdots & \mathcal{F}_L(\mathbf{b}_M^{(l)})\end{pmatrix},$$

$$\mathcal{F}_L(\mathbf{b}_i^{(j)})\triangleq\begin{pmatrix}b_i^{(j)}(0) & \cdots & b_i^{(j)}(L) & & \\ & \ddots & & \ddots & \\ & & b_i^{(j)}(0) & \cdots & b_i^{(j)}(L)\end{pmatrix}. \tag{4}$$

2) Data Selection Transform (DST) [2]: We have (5), which appears at the bottom of the page, where $\mathbf{B}^*$ is the complex conjugate of $\mathbf{B}$, $\otimes$ denotes the Kronecker product,

$$\mathbf{T} = [\mathbf{T}_{12}, \cdots, \mathbf{T}_{1M}, \cdots, \mathbf{T}_{(M-1)M}] \otimes \mathbf{I}_{L+1}$$

and $\mathbf{T}_{ij}$, $(i < j)$ is a $M \times M$ matrix whose (i, j)th entry is 1 and (j, i)th entry is -1 and is zero elsewhere. $\bar{\mathbf{I}}$ is an $[M(M-1)/2] \times [M(M-1)/2]$ identity matrix.

The following properties are direct consequences of the definitions. Let $\mathbf{A} \in C^{M(L+1)\times l}$, $\mathbf{B} \in C^{M(L+1)\times k}$, and $\alpha, \beta \in C^1$, and then

Symmetry:

$$\mathbf{A}^H \mathcal{H}_{M,L}(\mathbf{B}) = \mathbf{0} \Longleftrightarrow \mathbf{B}^H \mathcal{H}_{M,L}(\mathbf{A}) = \mathbf{0} \tag{7}$$

$$\mathbf{A}^H \mathcal{D}_{M,L}(\mathbf{B}) = \mathbf{0} \Longleftrightarrow \mathbf{B}^H \mathcal{D}_{M,L}(\mathbf{A}) = \mathbf{0}. \tag{8}$$

Orthogonality:

$$\mathcal{D}_{M,L}(\mathbf{h})^H \mathcal{H}_{M,L}(\mathbf{h}) = \mathbf{0}. \tag{9}$$

Commutativity:

$$\mathcal{H}_{M,L}(\mathcal{D}_{M,L}(\mathbf{h})) = \mathcal{D}_{M,L}(\mathcal{H}_{M,L}(\mathbf{h})). \tag{10}$$

Remarks:

1) The symmetry property is based on the commutativity of convolution. Equation (7) was also given in [3] and [5].
2) Orthogonality and commutativity between MFT and DST are special features of SIMO system, which are key properties used in this paper.
3) We note that in general, the columns of $\mathcal{D}_{M,L}(\mathbf{h})$ do not provide the entire null space. The construction of the orthogonal complement of the channel matrix is given in [1].

III. A Connection Between the LS and the SS Approaches

Although the SS and LS approaches are motivated differently, they can be derived in the same framework using properties of MFT and DST. We examine their connections in both identification and estimation aspects.

A. Identification Equations

The covariance matrix of received data is given by

$$\begin{aligned}\mathbf{R}_y(\mathbf{h}) &\triangleq E\{\mathbf{y}(t)\mathbf{y}^H(t)\} \\ &= \mathcal{H}_{M,L}(\mathbf{h})\mathbf{R}_w \mathcal{H}_{M,L}(\mathbf{h})^H + \sigma^2 \mathbf{I}.\end{aligned} \tag{11}$$

Let the singular value decomposition (SVD) of $\mathbf{R}_y(\mathbf{h})$ be

$$\begin{aligned}\mathbf{R}_y(\mathbf{h}) &= \mathbf{U}(\mathbf{h})\Lambda(\mathbf{h})\mathbf{U}^H(\mathbf{h}) \\ &= \mathbf{U}_s(\mathbf{h})\Lambda_s(\mathbf{h})\mathbf{U}_s^H(\mathbf{h}) + \sigma^2 \mathbf{U}_n(\mathbf{h})\mathbf{U}_n^H(\mathbf{h})\end{aligned} \tag{12}$$

where $\mathbf{U}_s(\mathbf{h})$ consists of the singular vectors associated with singular values greater than σ^2, and $\mathbf{U}_n(\mathbf{h})$ consists of the singular vectors associated with singular values equal to σ^2. Columns of $\mathbf{U}_s(\mathbf{h})$ and $\mathbf{U}_n(\mathbf{h})$ span the "signal" and "noise" subspaces, respectively. The following theorem gives the identification equations for LS and SS methods and shows the uniqueness of their identification.

Theorem 1: Given $\mathbf{R}_y(\mathbf{h})$ and its noise eigenmatrix $\mathbf{U}_n(\mathbf{h})$, the following relationships (13)–(15) hold and are as well equivalent.

a) $\tilde{\mathbf{h}} = \alpha\mathbf{h}, \quad \forall \alpha \in C^1 - \{0\} \Leftrightarrow$
$$\tilde{h}_i(z)h_j(z) = \tilde{h}_j(z)h_i(z), \quad \forall i, j. \tag{13}$$

b) $\tilde{\mathbf{h}}^H \mathbf{T}\{\bar{\mathbf{I}} \otimes [\mathbf{R}_y^*(\mathbf{h}) - \sigma^2\mathbf{I}]\}\mathbf{T}^H \tilde{\mathbf{h}} = 0 \Leftrightarrow$
$$\mathcal{D}_{M,L}(\mathbf{h})^H \mathcal{H}_{M,L}(\tilde{\mathbf{h}}) = \mathbf{0} \tag{14}$$

c) $\tilde{\mathbf{h}}^H \mathcal{H}_{M,L}(\mathbf{U}_n(\mathbf{h})) = 0 \Leftrightarrow$
$$\mathbf{U}_n^H(\mathbf{h})\mathcal{H}_{M,L}(\tilde{\mathbf{h}})\mathbf{0} \tag{15}$$

where $\tilde{h}_i(z)$ and $h_i(z)$ are z-transforms of $\tilde{h}_i(n)$ and $h_i(n)$, respectively.

Remark: Equations (14) and (15) in b) and c) provide identification equations of the LS and the SS methods, respectively. The equivalence of a), b), and c) shows that the identification equations of LS and SS have a unique solution. Differences between the LS and SS approaches are also evident from (14) and (15). The SS method uses the entire noise subspace spanned by the columns of $\mathbf{U}_n(\mathbf{h})$, whereas the LS method uses only the noise subspace spanned by the columns of $\mathcal{D}_{M,L}(\mathbf{h})$. In particular, range$\{\mathcal{D}_{M,L}(\mathbf{h})\} \subseteq$ range$\{\mathbf{U}_n(\mathbf{h})\}$. The use of the entire noise subspace is not necessary. Moulines *et al.* also gave a variation of the SS approach by using a part of the noise subspace [3].

Proof: First, we prove the equivalences within a), b), and c). The equivalence in c) is an obvious consequence of the symmetry property of MFT (7). For a), it is obvious that $\tilde{\mathbf{h}} = \alpha\mathbf{h} \Rightarrow \tilde{h}_i(z)h_j(z) = \tilde{h}_j(z)h_i(z), \forall i, j$. To show the converse since $\mathcal{Z}[h_i(z)] \subseteq \mathcal{Z}[\tilde{h}_i(z)] \bigcup \mathcal{Z}[h_j(z)]$. Hence, under A1, $\mathcal{Z}[h_i(z)] \subseteq \bigcap_{j=1}^M \{\mathcal{Z}[\tilde{h}_i(z)] \cup \mathcal{Z}[h_j(z)]\} = \mathcal{Z}[\tilde{h}_i(z)] \cup \bigcap_{j=1}^M \{\mathcal{Z}[h_j(z)]\} = \mathcal{Z}[\tilde{h}_i(z)]$. Since $\deg\{\tilde{h}_i(z)\} \le L$, $\tilde{h}_i(z) = \alpha h_i(z)$ for an arbitrary constant α. For b)

$$\begin{aligned}&\tilde{\mathbf{h}}^H \mathbf{T}[\bar{\mathbf{I}} \otimes (\mathbf{R}_y^* - \sigma^2\mathbf{I})]\mathbf{T}^H \tilde{\mathbf{h}} \\ &= \tilde{\mathbf{h}}^H \mathbf{T}\{\bar{\mathbf{I}} \otimes [\mathcal{H}_{M,L}(\mathbf{h})^* \mathbf{R}_w^* \mathcal{H}_{M,L}(\mathbf{h})^T]\}\mathbf{T}^H \tilde{\mathbf{h}} \quad (16) \\ &= \tilde{\mathbf{h}}^H \mathbf{T}\{\bar{\mathbf{I}} \otimes [\mathcal{H}_{M,L}(\mathbf{h})^*]\}(\bar{\mathbf{I}} \otimes \mathbf{R}_w^*) \\ &\quad [\bar{\mathbf{I}} \otimes \mathcal{H}_{M,L}(\mathbf{h})^T]\mathbf{T}^H \tilde{\mathbf{h}} \quad (17) \\ &= \{\tilde{\mathbf{h}}^H \mathcal{D}_{M,L}(\mathcal{H}_{M,L}(\mathbf{h}))\}(\bar{\mathbf{I}} \otimes \mathbf{R}_w^*) \\ &\quad \{\tilde{\mathbf{h}}^H \mathcal{D}_{M,L}(\mathcal{H}_{M,L}(\mathbf{h}))\}^H. \quad (18)\end{aligned}$$

Since $\mathbf{R}_w > 0$, $\tilde{\mathbf{h}}^H \mathbf{T}[\bar{\mathbf{I}} \otimes (\mathbf{R}_y^* - \sigma^2\mathbf{I})]\mathbf{T}^H \tilde{\mathbf{h}} = 0$ if and only if $\tilde{\mathbf{h}}^H \mathcal{D}_{M,L}(\mathcal{H}_{M,L}(\mathbf{h})) = 0$. By the commutativity property (10), $\tilde{\mathbf{h}}^H \mathcal{D}_{M,L}(\mathcal{H}_{M,L}(\mathbf{h})) = \tilde{\mathbf{h}}^H \mathcal{H}_{M,L}(\mathcal{D}_{M,L}(\mathbf{h})) = 0$. By the symmetry property (7), $\mathcal{D}_{M,L}(\mathbf{h})^H \mathcal{H}_{M,L}(\tilde{\mathbf{h}}) = 0$. Now, we prove that a) $\Rightarrow$ c) $\Rightarrow$ b) $\Rightarrow$ a).

1) a) $\Rightarrow$ c): Since $\mathbf{U}_n^H(\mathbf{h})\mathcal{H}_{M,L}(\mathbf{h}) = 0$, therefore, $\mathbf{U}_n^H(\mathbf{h})\mathcal{H}_{M,L}(\tilde{\mathbf{h}}) = \alpha \mathbf{U}_n^H(\mathbf{h})\mathcal{H}_{M,L}(\mathbf{h}) = \mathbf{0}$.
2) c) $\Rightarrow$ b): From the orthogonality property, $\mathcal{D}_{M,L}(\mathbf{h})^H \mathcal{H}_{M,L}(\mathbf{h}) = 0$. From (15), there is an $\mathbf{A}$ such that $\mathcal{D}_{M,L}(\mathbf{h}) = \mathbf{U}_n(\mathbf{h})\mathbf{A}$. We now have $\mathcal{D}_{M,L}(\mathbf{h})^H \mathcal{H}_{M,L}(\tilde{\mathbf{h}}) = \mathbf{A}^H \mathbf{U}_n^H(\mathbf{h})\mathcal{H}_{M,L}(\tilde{\mathbf{h}}) = \mathbf{0}$.
3) b) $\Rightarrow$ a): $\mathcal{D}_{M,L}(\mathbf{h})^H \mathcal{H}_{M,L}(\tilde{\mathbf{h}}) = 0$ is the matrix form of $\tilde{h}_i(z)h_j(z) = \tilde{h}_j(z)h_i(z), \forall i, j$. □

$$\mathcal{D}_{M,L}(\mathbf{B}) \triangleq \begin{pmatrix} \mathbf{B}_2 & \mathbf{B}_3 & \cdots & \mathbf{B}_M & \vdots & & & & \vdots & \vdots & \\ -\mathbf{B}_1 & & & & \vdots & \mathbf{B}_3 & \cdots & \mathbf{B}_M & \vdots & \vdots & \\ & -\mathbf{B}_1 & & & \vdots & -\mathbf{B}_2 & & & \vdots & \cdots & \vdots & \\ & & \ddots & & \vdots & & \ddots & & \vdots & \vdots & \mathbf{B}_M \\ & & & -\mathbf{B}_1 & \vdots & & & -\mathbf{B}_2 & \vdots & \vdots & -\mathbf{B}_{M-1} \end{pmatrix}^* = \mathbf{T}(\bar{\mathbf{I}} \otimes \mathbf{B}^*) \tag{5}$$

B. The LS and SS Estimators

In channel estimation, $\mathbf{R}_y(\mathbf{h})$ may be replaced by the sample covariance $\hat{\mathbf{R}}_y$. There is a number of different implementations of the LS and SS methods (see e.g., [2] and [3]). In this correspondence, we shall consider the LS and SS estimators defined by

$$\hat{\mathbf{h}}_{LS} = \arg\min_{||\hat{\mathbf{h}}||=1} \hat{\mathbf{h}}^H[\mathbf{T}(\bar{\mathbf{I}} \otimes \hat{\mathbf{R}}_y^*)\mathbf{T}^H]\hat{\mathbf{h}}. \quad (19)$$

$$\hat{\mathbf{h}}_{SS} = \arg\min_{||\hat{\mathbf{h}}||=1} \hat{\mathbf{h}}^H \mathcal{H}_{M,L}(\hat{\mathbf{U}}_n)\mathcal{H}_{M,L}(\hat{\mathbf{U}}_n)^H \hat{\mathbf{h}}. \quad (20)$$

In the original development [2], the LS estimator was referred to as a *deterministic* approach that minimizes the following least-squares cost: $\min_{||\hat{\mathbf{h}}||=1} ||\mathcal{D}_{M,L}([\mathbf{x}(t)])^H\hat{\mathbf{h}}||^2$. Since $\mathcal{D}_{M,L}([\mathbf{x}(t)])\mathcal{D}_{M,L}([\mathbf{x}(t)])^H = \mathbf{T}(\bar{\mathbf{I}} \otimes \hat{\mathbf{R}}_y^*)\mathbf{T}^H$, it is easy to verify that this optimization is equivalent to (19). The following theorem shows that when $M = 2$, the two estimators are identical.

Theorem 2: When $M = 2$, $\hat{\mathbf{h}}_{LS} = \hat{\mathbf{h}}_{SS}$ with probability one.

Proof: For this case, $\mathcal{D}_{2,L}(\mathbf{h}) = \mathbf{T}\mathbf{h}^*$, and $\mathbf{T} = [\mathbf{T}_{12}] \otimes \mathbf{I}_{L+1} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{L+1} \\ -\mathbf{I}_{L+1} & \mathbf{0} \end{pmatrix}$ is an orthogonal matrix. Assumption A1 implies that (see [6]) rank$[\mathcal{H}_{2,L}(\mathbf{h})] = 2L+1$. The dimension of the noise subspace of $\mathbf{U}_n(\mathbf{h})$ is 1. Therefore, range$\{\mathbf{U}_n(\mathbf{h})\}$ = range$\{\mathcal{D}_{2,L}(\mathbf{h})\}$ = range$\{\mathbf{T}\mathbf{h}^*\}$. When the sample covariance $\hat{\mathbf{R}}_y$ is used, the estimated noise subspace is given by the eigenvector $\hat{\mathbf{g}}$ associated with the smallest eigenvalue. Note also that with probability 1, $\hat{\mathbf{g}}$ satisfies assumption A1. Hence $\mathcal{H}_{2,L}(\hat{\mathbf{g}})$ is row-rank deficient by 1. Again, by the orthogonality property, $\mathcal{D}_{2,L}(\hat{\mathbf{g}})^H\mathcal{H}_{2,L}(\hat{\mathbf{g}}) = \mathbf{0}$. Therefore

$$\begin{aligned}\hat{\mathbf{h}}_{SS} &= \arg\min_{||\hat{\mathbf{h}}||=1} \hat{\mathbf{h}}^H\mathcal{H}_{2,L}(\hat{\mathbf{g}})\mathcal{H}_{2,L}(\hat{\mathbf{g}})^H\hat{\mathbf{h}} \\ &= \mathcal{D}_{2,L}(\hat{\mathbf{g}}) = \mathbf{T}\hat{\mathbf{g}}^*. \end{aligned} \quad (21)$$

On the other hand, since $\mathbf{T}$ is orthogonal

$$\begin{aligned}\hat{\mathbf{h}}_{LS} &= \arg\min_{||\hat{\mathbf{h}}||=1} \hat{\mathbf{h}}^H(\mathbf{T}\hat{\mathbf{R}}_y^*\mathbf{T}^H)\hat{\mathbf{h}} \\ &= \arg\min_{||\hat{\mathbf{h}}||=1} (\mathbf{T}^H\hat{\mathbf{h}})^H\hat{\mathbf{R}}_y^*(\mathbf{T}^H\hat{\mathbf{h}}) \\ &= (\mathbf{T}^H)^{-1} \arg\min_{||\hat{\mathbf{h}}||=1} \hat{\mathbf{h}}^H\hat{\mathbf{R}}_y^*\hat{\mathbf{h}} = \mathbf{T}\hat{\mathbf{g}}^*. \end{aligned} \quad (22)$$

We now have shown that the two estimators are identical. □

We note that the above theorem is not true when $M > 2$. Such a case may be the result of using a sampling rate higher than twice the symbol rate or when multiple receivers are involved. The difference between the two algorithms is manifested in the ways the noise subspaces are used. The SS approach uses $\mathbf{U}_n(\mathbf{h})$, which is the entire noise subspace. The LS approach uses a special noise subspace $\mathcal{D}_{M,L}(\mathbf{h})$.

IV. Conclusion

By exploiting key properties of the multichannel filtering and the data selection transforms, we gave a unified presentation of identification and estimation using LS and SS approaches. We also show that the two estimators are identical for the special case that there are two subchannels. The differences between the two approaches are revealed in their utilizations of the noise subspace, especially when more that two subchannels are involved.

Acknowledgment

The authors wish to acknowledge comments from Prof. Willett that improved the presentation of this paper.

References

[1] Y. Hua, H. Yang, and M. Zhou, "Blind system identification using multiple sensors," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing,* Detroit, MI, May 1995, vol. 5, pp. 3171–3174.

[2] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind identification of multichannel FIR systems," in *Proc. 27th Asilomar Conf. Signal, Syst., Comp.,* Asilomar, CA, Oct. 1993.

[3] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace-methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing,* vol. 43, no. 2, pp. 516–525, Feb. 1995.

[4] S. V. Schell, D. L. Smith, and S. Roy, "Blind channel identification using subchannel response matching," in *Proc. 26th Conf. Information Sciences and Systems,* Princeton, NJ, Mar. 1994.

[5] D. Slock, "Blind fractionally-spaced equalization, perfect reconstruction filterbanks, and multilinear prediction," in *Proc. ICASSP'94 Conf.,* Adelaide, Australia, Apr. 1994.

[6] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind identification and equalization of multipath channels: A frequency domain approach," *IEEE Trans. Inform. Theory,* vol. 41, no. 1, pp. 329–334, Jan. 1995.

[7] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conf.,* Pacific Grove, CA, Nov. 1991.

[8] ——, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory,* vol. 40, no. 2, Mar. 1994.

Alignment Blur in Coherently Averaged Images

D. M. Monro and D. M. Simpson

Abstract—**Blurring of coherently averaged images due to imperfect alignment is studied, and two restoration methods are proposed and evaluated. It is shown that iterative realignment is more powerful than post-filtering in reducing blur. The value of averaging and restoration is illustrated on human subjects in noisy video sequences.**

I. Introduction

In signal and image processing applications, data is often corrupted by noise, and many techniques have been proposed to reduce its effect. If multiple aligned versions of the data are available, each with uncorrelated additive noise, signal averaging will increase the signal-to-noise-ratio (SNR). This technique is widely applied in signal processing but is used less often in the enhancement of images, perhaps because of the additional degrees of freedom arising from the

Manuscript received July 20, 1991; revised December 19, 1995. This work, and related work on moving image registration and reconstruction, were supported, in part, by the UK Science and Engineering Research Council, Grant no. GR/D92370. The associate editor coordinating the review of this paper and approving it for publication was Prof. Aggelos K. Katsaggelos.

D. M. Monro is with the School of Electrical Engineering, University of Bath, Claverton Down, Bath, BA2 7AY, England.

D. M. Simpson is with the School of Electrical Engineering, University of Bath, Claverton Down, Bath, BA2 7AY, England and the Biomedical Engineering Program, University of Rio de Janeiro, OPPE/UFRJ, P.O. Box 68510, 21945-970 Rio de Janeiro, Brazil.

Publisher Item Identifier S 1053-587X(96)03944-X.