

- [5] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
- [6] D. S. Burdick, Private communication.
- [7] D. D. Feldman and L. J. Griffiths, "A constraint projection approach to robust adaptive beamforming," in *Proc. IEEE ICASSP-88*, 1988, pp. 1381-1384.
- [8] J. L. Hunt, "Statistics relevant to the detection performance of adaptive beamformers," Progress Rep. to R.A.E. under contract no. AT/2040/0183 A.U.W.W., 1982.
- [9] P. J. Bickel and K. A. Doksum, *Mathematical Statistics*. San Francisco: Holden-Day, 1977.

## Large Modular Structures for Adaptive Beamforming and the Gram-Schmidt Preprocessor

Rajesh Sharma and Barry D. Van Veen

**Abstract**—An alternate derivation of the modular structure for linearly constrained minimum variance beamforming proposed in [1] is presented using a vector space approach. This approach eliminates the tedious algebra employed in [1] and establishes the relationship between the modular structure and the Gram-Schmidt preprocessor [3]. The modular structure is obtained using a factorization of the orthogonal projection operator in Hilbert space. The Gram-Schmidt preprocessor is a special case of the general modular decomposition. It is also shown that these structures offer computational efficiencies when multiple beamformers are implemented simultaneously.

### I. INTRODUCTION

**I**N linearly constrained minimum variance (LCMV) beamforming, the beamformer weights are constrained by a set of linear equations [4]. The constraints are used to control the beamformer response over specified directions and frequencies. The weights are chosen to minimize output variance while satisfying the constraint equations. The generalized sidelobe canceller (GSC) is an implementation of the LCMV weight vector that is well suited for adaptive algorithms. In [1], it was shown that the GSC could be decomposed into a cascade of adaptive modules, and the advantages of this decomposition were discussed.

In this correspondence, we derive the equivalence between the GSC and the modular structure in a different manner. The problem is formulated as an optimization problem in a Hilbert space, which enables one to use the special properties of Hilbert spaces, and thus gain insight into the operation of the modular structure. A simple proposition is proved that establishes the factorization of the orthogonal projection operator analogous to the factorization used in the modified Gram-Schmidt algorithm [5]. This approach eliminates the use of tedious algebra in the derivation of the modular structure. It is shown that the modular structure performs an orthogonalization of the estimation subspace similar to that performed by the Gram-Schmidt preprocessor [3]. Furthermore, it is shown that the Gram-

Manuscript received November 3, 1992; revised June 25, 1993. This work was supported in part by the Army Research Office under Grant DAAL03-89-K-0141 and by the National Science Foundation under Award MIP-895 8559. The associate editor coordinating the review of this paper and approving it for publication was Prof. John A. Stuller.

The authors are with the University of Wisconsin-Madison, Madison, WI 53706.

IEEE Log Number 9214190.

Schmidt preprocessor is obtained as a special case of the modular decomposition and relations are established between the parameters of both structures.

The outline of this correspondence is as follows. In Section II, a brief introduction to LCMV beamforming is provided and the Hilbert space formulation of the LCMV problem is given. The derivation of the modular structure is given in Section III. Comparisons are made between the modular structure and the Gram-Schmidt preprocessor in Section IV. A discussion of the results is provided in Section V.

### II. LINEARLY CONSTRAINED MINIMUM VARIANCE BEAMFORMING

Let  $\mathbf{w}_{N \times 1}$  denote the beamformer weight vector. The beamformer output,  $y$ , due to the input,  $\mathbf{x} = [x_1, \dots, x_N]^T$  is

$$y = \mathbf{w}^H \mathbf{x}.$$

Here the superscript  $H$  denotes complex conjugate transpose, and as a matter of convention, vector and matrix quantities will be denoted by boldface symbols. If  $\mathbf{C}_{N \times L}$  and  $\mathbf{f}_{L \times 1}$  denote the constraint matrix and desired response vector respectively<sup>1</sup>, then the LCMV beamforming problem is

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \quad \text{such that} \quad \mathbf{C}^H \mathbf{w} = \mathbf{f} \quad (1)$$

where  $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H]$  is the data covariance matrix and  $E[\ast]$  is the expectation operation. The constraint set in (1) is a linear variety in the  $N$ -dimensional complex Euclidean space,  $C^N$ . If  $\mathbf{w}_q = \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} \mathbf{f}^2$  and the columns of an  $N \times (N - L)$  matrix  $\mathbf{C}_n$  form a basis for the orthogonal complement of the space spanned by the columns of  $\mathbf{C}$ , then the constraint set  $\Gamma$  is given by

$$\Gamma = \{\mathbf{w} \in C^N : \mathbf{w} = \mathbf{w}_q - \mathbf{C}_n \mathbf{w}_n, \mathbf{w}_n \in C^{N-L}\}. \quad (2)$$

Since the number of columns in  $\mathbf{C}_n$  is  $N - L$ , we say that the beamformer has  $N - L$  adaptive degrees of freedom. Let  $P = N - L$ . The linear variety  $\Gamma$  defined in (2) depends only on the span of the columns of  $\mathbf{C}_n$  and is independent of the particular choice of  $\mathbf{w}_q$  so long as it satisfies the constraint equation.

If  $(\Omega, \mathcal{F}, \varphi)$  denotes an underlying probability space, then

$$L_2 = \{x : E[x] = 0, E[|x|^2] < \infty\}$$

is a Hilbert space [2], with the inner product defined as

$$(x|y) = E[y^* x] \quad \forall x, y \in L_2$$

and norm defined by

$$\|x\| = (x|x)^{\frac{1}{2}} \quad \forall x \in L_2.$$

Without loss of generality, we shall assume that all random variables in this paper are elements of the  $L_2$  Hilbert space. Assuming  $\mathbf{R}$  is positive definite, then  $\mathcal{H} = \text{span}\{\mathbf{x}\}$  is a  $N$ -dimensional Hilbert space<sup>3</sup>. Therefore, (1) is rewritten using (2) as

$$\min_{\mathbf{w}_n} \|\mathbf{w}_q^H \mathbf{x} - \mathbf{w}_n^H \mathbf{C}_n^H \mathbf{x}\|^2,$$

with the optimum beamformer weight vector  $\mathbf{w}$  given by  $\mathbf{w} = \mathbf{w}_q - \mathbf{C}_n \mathbf{w}_n$ . Hence, the beamformer output,  $y$  is

$$y = \mathbf{w}_q^H \mathbf{x} - \mathcal{P}(\mathbf{w}_q^H \mathbf{x} | \mathcal{M}), \quad (3)$$

<sup>1</sup> $L$  denotes the number of constraints ( $L < N$ ).

<sup>2</sup> $\mathbf{C}$  is assumed to be full rank.

<sup>3</sup>Here span of a random vector denotes the span of the random variables that are the elements of the vector.

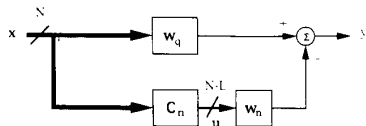


Fig. 1. The generalized sidelobe canceller.

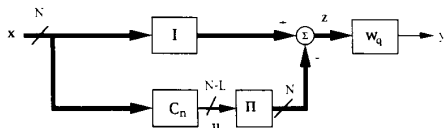


Fig. 2. Decoupling of the adaptive and nonadaptive components.

where  $\mathcal{M} = \text{span}\{\mathbf{u}\}$  and  $\mathbf{u} = [u_1, \dots, u_{N-L}]^T = \mathbf{C}_n^H \mathbf{x}$ .  $\mathcal{P}(\mathbf{w}_q^H \mathbf{x} | \mathcal{M})$  denotes the projection of the random variable  $\mathbf{w}_q^H \mathbf{x}$  onto the subspace  $\mathcal{M}$ . The subspace  $\mathcal{M}$  is called the estimation subspace or the adaptation space. The number of adaptive degrees of freedom is the dimension of the estimation subspace. The estimation subspace depends on the span of the columns of  $\mathbf{C}_n$ , not the specific  $\mathbf{C}_n$  utilized. It is easily shown that  $\mathcal{P}$  is a linear operator. Application of the orthogonality principle [2] yields  $\mathbf{w}_n = (\mathbf{C}_n^H \mathbf{R} \mathbf{C}_n)^{-1} \mathbf{C}_n^H \mathbf{R} \mathbf{w}_q$ .

### III. MODULAR DECOMPOSITION

In block diagram form, (3) yields the GSC structure of Fig. 1. Using the linearity of the projection operator, (3) is rewritten as

$$y = \mathbf{w}_q^H (\mathbf{x} - \mathcal{P}(\mathbf{x} | \mathcal{M})). \quad (4)$$

Here,  $\mathcal{P}(\mathbf{x} | \mathcal{M}) = [\mathcal{P}(x_1 | \mathcal{M}), \dots, \mathcal{P}(x_N | \mathcal{M})]^T$ . Equation (4) suggests the structure shown in Fig. 2. The  $i^{\text{th}}$  column of the matrix  $\mathbf{\Pi}_{(N-L) \times N}$  consists of the optimum coefficients used to predict  $x_i$  by a linear combination of the  $\{u_i\}_{i=1}^P$ . Application of the orthogonality principle indicates that  $E[\mathbf{C}_n^H \mathbf{x} \mathbf{x}^H (\mathbf{I} - \mathbf{C}_n \mathbf{\Pi})] = \mathbf{0}$ , yielding

$$\mathbf{\Pi} = (\mathbf{C}_n^H \mathbf{R} \mathbf{C}_n)^{-1} \mathbf{C}_n^H \mathbf{R}.$$

In Fig. 2, the elements of the random vector  $\mathbf{z} = [z_1, \dots, z_N]^T$  are the error random variables, that is  $\mathbf{z} = \mathbf{x} - \mathcal{P}(\mathbf{x} | \mathcal{M})$ .

**Definition:** A subspace  $\mathcal{M}$  is a direct sum of two subspaces  $\mathcal{M}_1$  and  $\mathcal{M}_2$  denoted by  $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$  if every  $m \in \mathcal{M}$  has a unique representation of the form  $m = m_1 + m_2$  with  $m_1 \in \mathcal{M}_1$  and  $m_2 \in \mathcal{M}_2$ .

**Definition:** Two subspaces  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are said to be orthogonal if for all  $m_1 \in \mathcal{M}_1$  and  $m_2 \in \mathcal{M}_2$  we have  $(m_1 | m_2) = 0$ .

Unless otherwise stated we assume that the subspaces which form the direct sum of a subspace are mutually orthogonal or equivalently, that all direct decompositions are orthogonal direct decompositions (i.e., if  $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$  then  $\mathcal{M}_1 \perp \mathcal{M}_2$ ).

**Proposition:** Let  $\mathcal{M} \subset \mathcal{H}$  be a  $P$  dimensional subspace and  $d \in \mathcal{H}$ . If  $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$  then

$$d - \mathcal{P}(d | \mathcal{M}) = d - \mathcal{P}(d | \mathcal{M}_1) - \mathcal{P}([d - \mathcal{P}(d | \mathcal{M}_1)] | \mathcal{M}_2). \quad (5)$$

**Proof:**  $d - \mathcal{P}(d | \mathcal{M}) = d - \mathcal{P}(d | \mathcal{M}_1) - \mathcal{P}(d | \mathcal{M}_2)$ , since  $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$ . Now  $\mathcal{P}(\mathcal{P}(d | \mathcal{M}_1) | \mathcal{M}_2) = 0$  because  $\mathcal{M}_1 \perp \mathcal{M}_2$  and  $\mathcal{P}(d | \mathcal{M}_1) \in \mathcal{M}_1$ . Thus,  $d - \mathcal{P}(d | \mathcal{M}) = d - \mathcal{P}(d | \mathcal{M}_1) - \mathcal{P}(d | \mathcal{M}_2) + \mathcal{P}(\mathcal{P}(d | \mathcal{M}_1) | \mathcal{M}_2)$  from which (5) is obtained using the linearity of the projection operator.  $\square$

To simplify notation, let

$$\mathcal{P}_\perp(d | \mathcal{M}) = d - \mathcal{P}(d | \mathcal{M}).$$

That is,  $\mathcal{P}_\perp$  denotes the orthogonal projection operator. Alternatively,  $\mathcal{P}_\perp(d | \mathcal{M}) = \mathcal{P}(d | \mathcal{M}^\perp)$ , where  $\mathcal{M}^\perp = \{x \in \mathcal{H} : (x | y) = 0 \forall y \in \mathcal{M}\}$  and  $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ . Therefore, (4) and (5) are written as

$$y = \mathbf{w}_q^H \mathcal{P}_\perp(\mathbf{x} | \mathcal{M}) \quad (6)$$

$$\mathcal{P}_\perp(d | \mathcal{M}) = \mathcal{P}_\perp(\mathcal{P}_\perp(d | \mathcal{M}_1) | \mathcal{M}_2). \quad (7)$$

If  $\mathcal{M} = \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_Q$ , then (7) implies

$$\mathcal{P}_\perp(d | \mathcal{M}) = \mathcal{P}_\perp(\dots \mathcal{P}_\perp(\mathcal{P}_\perp(d | \mathcal{M}_1) | \mathcal{M}_2) | \dots | \mathcal{M}_Q). \quad (8)$$

The factorization of the orthogonal projection operator in (8) is analogous to the factorization of the  $\mathcal{P}_\perp$  operator in Euclidean space in the Modified Gram-Schmidt (MGS) algorithm [5]. This is illustrated by the following example.

**Example:** As shown in [5], the primary difference in performing orthogonalization via the Classical Gram-Schmidt (CGS) algorithm and the Modified Gram-Schmidt (MGS) algorithm in complex Euclidean space  $C^N$  is the implementation of the orthogonal projection operator at each step. If at step  $i+1$  the orthonormal set of vectors are the column vectors of the  $N \times i$  matrix  $\mathbf{Q}_i = [\mathbf{q}_1, \dots, \mathbf{q}_i]$ , then the orthogonal projection operator at step  $i+1$  is given by the  $N \times N$  matrix  $(\mathbf{I} - \mathbf{Q}_i \mathbf{Q}_i^H)$ . This orthogonal projection operator is factored as

$$(\mathbf{I} - \mathbf{Q}_i \mathbf{Q}_i^H) = (\mathbf{I} - \mathbf{q}_1 \mathbf{q}_1^H) \dots (\mathbf{I} - \mathbf{q}_i \mathbf{q}_i^H). \quad (9)$$

In CGS the  $\mathcal{P}_\perp$  operator is implemented directly as  $\mathbf{I} - \mathbf{Q}_i \mathbf{Q}_i^H$ , however in the MGS the  $\mathcal{P}_\perp$  operator is implemented in the factored form as a sequence of  $i$  linear transformations given in (9). This is analogous to (8) with the only difference being that the subspaces  $\mathcal{M}_j$  need not be one dimensional. However, the factorization in (9) is valid even if the  $\mathbf{q}_j$  column vectors are matrices, provided the columns of  $\mathbf{q}_j$  are orthonormal and span mutually orthogonal subspaces. In this case the factorization in (9) is the Euclidean version of (8).  $\square$

The modular structure is now derived using (8). Since  $\mathbf{z} = \mathcal{P}_\perp(\mathbf{x} | \mathcal{M})$ , if  $\mathcal{M} = \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_Q$ , then

$$\mathbf{z} = \mathcal{P}_\perp(\dots \mathcal{P}_\perp(\mathcal{P}_\perp(\mathbf{x} | \mathcal{M}_1) | \mathcal{M}_2) | \dots | \mathcal{M}_Q). \quad (10)$$

Equation (10) establishes a procedure for computing  $\mathbf{z}$  in  $Q$  stages provided the random variables that span  $\mathcal{M}_j$  are available at stage  $j$ . That is, if  $\mathbf{z}_i = \mathcal{P}_\perp(\mathbf{z}_{i-1} | \mathcal{M}_i)$  for  $i = 1, \dots, Q$  with  $\mathbf{z}_0 = \mathbf{x}$  then  $\mathbf{z}_Q = \mathbf{z}$ . To obtain the orthogonal direct decomposition  $\mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_Q$  of the estimation subspace  $\mathcal{M}$  a sequence of  $Q$  successive orthogonalizations are utilized, analogous to those employed in the Gram-Schmidt algorithm. Let  $\mathbf{C}_n = [\mathbf{C}_1, \dots, \mathbf{C}_Q]$ , where  $\mathbf{C}_i$  has  $P_i$  columns and let  $\mathbf{u}_i = \mathbf{C}_i^H \mathbf{x}$  for  $i = 1, \dots, Q$ , where  $U_i = \text{span}\{\mathbf{u}_i\}$ . In general  $\{U_i\}_{i=1}^Q$  do not provide an orthogonal direct decomposition of  $\mathcal{M}$ . Therefore, to obtain an orthogonal direct decomposition of  $\mathcal{M}$ , define

$$\mathbf{v}_i = \mathbf{u}_i - \mathcal{P}(\mathbf{u}_i | \mathcal{M}_0 \oplus \dots \oplus \mathcal{M}_{i-1}), \quad i = 1, \dots, Q \quad (11)$$

where  $\mathcal{M}_0$  is the span of the zero random variable and  $\mathcal{M}_i = \text{span}\{\mathbf{v}_i\}$ . Equation (11) yields the orthogonal direct decomposition  $\mathcal{M} = \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_Q$ . Application of (8) to (11) yields

$$\mathbf{v}_i = \mathcal{P}_\perp(\dots \mathcal{P}_\perp(\mathcal{P}_\perp(\mathbf{u}_i | \mathcal{M}_0) | \mathcal{M}_1) | \dots | \mathcal{M}_{i-1}). \quad (12)$$

Using the linearity of the projection operator (12) is equivalently written as

$$\mathbf{v}_i = \mathbf{C}_i^H \mathcal{P}_\perp(\dots \mathcal{P}_\perp(\mathcal{P}_\perp(\mathbf{x} | \mathcal{M}_0) | \mathcal{M}_1) | \dots | \mathcal{M}_{i-1}). \quad (13)$$

Since  $\mathcal{P}_\perp(\mathbf{x} | \mathcal{M}_0) = \mathbf{z}_0$  (13) is rewritten as

$$\mathbf{v}_i = \mathbf{C}_i^H \mathbf{z}_{i-1} \quad i = 1, \dots, Q. \quad (14)$$

Equations (10) and (14) are combined to obtain the modular structure shown in Fig. 3. Each stage of the structure performs an orthogonal projection as described in (10), that is  $\mathcal{P}_\perp(\mathbf{z}_{i-1} | \mathcal{M}_i) = (\mathbf{I} - \mathbf{C}_i \mathbf{\Pi}_i)^H \mathbf{z}_{i-1}$ .

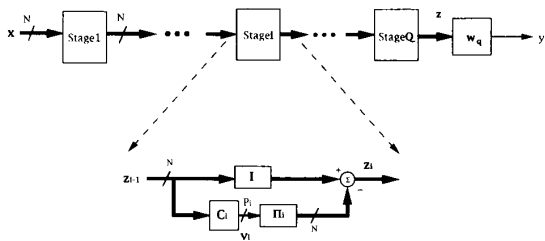


Fig. 3. The general modular decomposition.

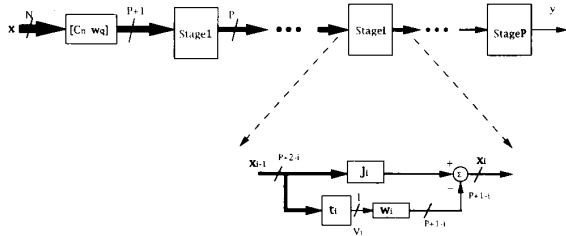


Fig. 4. The Gram-Schmidt preprocessor.

#### IV. GRAM-SCHMIDT PREPROCESSOR AND THE MODULAR STRUCTURE

The set of  $Q$  equations described in (11) is analogous to the Gram-Schmidt algorithm and is identical if the submatrices  $C_i$  are column vectors, in which case the  $u_i$ 's are scalar random variables. It follows that if  $C_i$  are column vectors, the modular structure orthogonalizes the set  $\{u_1, \dots, u_P\}$  via the MGS<sup>4</sup> algorithm to yield the orthogonal set  $\{v_1, \dots, v_P\}$ . This is precisely what is done in the Gram-Schmidt preprocessor [3]. The Gram-Schmidt preprocessor is a structure that solves (3) via the MGS algorithm and is most similar in operation to the modular structure shown in Fig. 3 when the  $C_i$  are column vectors. The Gram-Schmidt preprocessor orthogonalizes  $\{u_1, \dots, u_P, \mathbf{w}_q^H \mathbf{x}\}$  as follows:

$$v_i = \mathcal{P}_\perp(u_i | \mathcal{M}_0 \oplus \dots \oplus \mathcal{M}_{i-1}) \quad i = 1, \dots, (P+1), \quad (15)$$

where for notational convenience  $u_{P+1} = \mathbf{w}_q^H \mathbf{x}$  and as before  $\mathcal{M}_i = \text{span}\{v_i\}$  with  $\mathcal{M}_0$  the span of the zero random variable. It then follows from (15) that  $v_{P+1} = y$ . If the orthogonal projection operator in (15) is implemented using (8), then the Gram-Schmidt preprocessor is obtained, that is

$$v_i = \mathcal{P}_\perp(\dots \mathcal{P}_\perp(\mathcal{P}_\perp(u_i | \mathcal{M}_0) | \mathcal{M}_1) | \dots | \mathcal{M}_{i-1}). \quad (16)$$

The Gram-Schmidt preprocessor is shown in Fig. 4, where  $\mathbf{J}_i^H = [0_{P+1-i \times 1}, \mathbf{I}_{P+1-i}]$  and  $\mathbf{t}_i$  is a  $P+2-i$  column vector with a one in the  $i$ th position and zeros everywhere else. At the  $i$ th stage,  $v_i$  is computed and all the projections into  $\mathcal{M}_i$  are performed, and hence the computation described in (16) is implemented in a parallel manner.

The primary distinction between the modular structure and the Gram-Schmidt preprocessor is the transition from (12) to (13). Hence, the difference between the two structures is a direct result of the linearity of the projection operator. To illustrate this further, consider the equivalent GSC representation shown in Fig. 5. This representation follows directly from (3). Here,  $\mathbf{e} = [0, \dots, 1]^T$  and  $\mathbf{T}^H = [\mathbf{I}_P, \mathbf{0}_{P \times 1}]$ . Applying the modular decomposition to the structure in the dashed box (i.e.,  $\mathbf{T} \sim \mathbf{C}_n$  and  $\mathbf{e} \sim \mathbf{w}_q$ ) with  $\mathbf{T}$  partitioned into single columns yields a structure of the same form as that represented in Fig. 3. Unlike the general case, this structure

<sup>4</sup>Since the orthogonal projection operator of (11) is factored using (8).

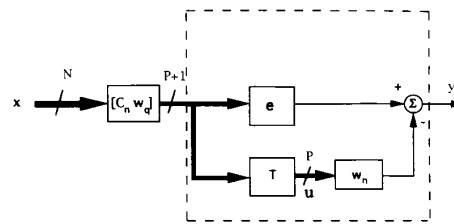


Fig. 5. Equivalent representation of the GSC.

can be further simplified. In the first stage the random variable  $u_1$  is used to predict itself which implies the first element of  $\mathbf{\Pi}_1$  is 1 and the first element of the random vector  $\mathbf{z}_1$  is 0. Therefore, a reduction in the dimension of the input to the second stage is possible. Repeating similar arguments for the remaining stages yields the Gram-Schmidt preprocessor shown in Fig. 4. The random vector outputs and filter parameters of the structures in Fig. 3 (assuming the  $C_i$  are column vectors) and Fig. 4 are related as follows: if  $\mathbf{A}_i = [C_{i+1}, \dots, C_P, \mathbf{w}_q]$ , then

$$\mathbf{x}_i = (\mathbf{A}_i)^H \mathbf{z}_i \quad i = 1, \dots, P.$$

$$\mathbf{w}_i = \mathbf{\Pi}_i \mathbf{A}_i \quad i = 1, \dots, P.$$

#### V. DISCUSSION

Often in practice the mathematical expectation used in the previous analysis is replaced by a sample mean. For instance, if we have available  $M$  different samples of the beamformer input denoted by  $\{\mathbf{x}[n]\}_{n=1}^M$ , where  $M > N$ , then the sample covariance matrix is given by  $\mathbf{R} = (1/M) \sum_{n=1}^M \mathbf{x}[n] \mathbf{x}^H[n]$ , and the sample variance (output power) of the beamformer is given by  $\mathbf{w}^H \mathbf{R} \mathbf{w}$ . Minor modifications to the previous analysis are required to accommodate the replacement of the mathematical expectation with the sample mean. The same notation is used in order to draw analogies with the previous work. Let  $\mathbf{X} = [\mathbf{x}[1], \dots, \mathbf{x}[M]]$  be a  $N \times M$  matrix consisting of the beamformer inputs and the  $1 \times M$  row vector  $\mathbf{y} = \mathbf{w}^H \mathbf{X}$  consist of the beamformer outputs. The previous analysis is now repeated noting that  $L_2$  space is replaced by  $M$ -dimensional complex Euclidean space  $C^M$  with the standard inner product inducing the Euclidean 2-norm. The random variables of the previous analysis are now  $M$ -dimensional row vectors elements of  $C^M$ . For example, the random variables  $x_i$  in the previous analysis are now replaced by the row vectors  $\mathbf{x}_i$  where  $\mathbf{x}_i$  is the  $i$ th row of the  $N \times M$  matrix  $\mathbf{X}$  and the row vector  $\mathbf{y}$  plays a role analogous to the random variable  $y$  of the previous analysis. One then proceeds to show the equivalence between the structures of Fig. 1 and Fig. 2 and so on.

It follows that application of sample matrix inversion (SMI) adaptive algorithm yields identical outputs and convergence rates for all the structures discussed previously since it is based on replacing the mathematical expectation by the sample mean. Application of the recursive least squares (RLS) adaptive algorithm is equivalent for the GSC of Fig. 1 and the structure in Fig. 2 for the same reasons. Application of the least mean square (LMS) adaptive algorithm yields identical outputs and convergence rates for the GSC and the structure in Fig. 2 as is easily verified by the LMS update equations. In general, application of the LMS and RLS adaptive algorithms yields different convergence rates when applied to the structures of Figs. 1, 3, and 4. Use of the SMI adaptive algorithm with the Gram-Schmidt preprocessor corresponds to the well known method of solving the least squares problem by the MGS algorithm [5].

If the  $\mathbf{T}$  matrix of Fig. 5 is partitioned arbitrarily and then the modular decomposition is applied, a different structure than the Gram-Schmidt preprocessor is obtained. This structure may be

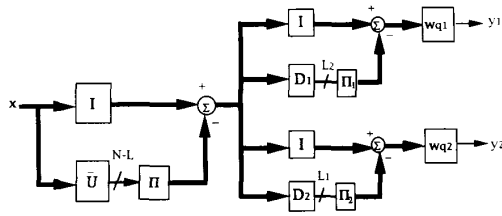


Fig. 6. Implementation of multiple beamformers.

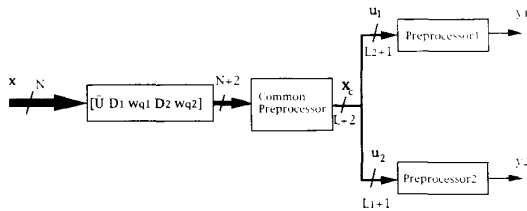


Fig. 7. Implementation of multiple beamformers via the Gram-Schmidt preprocessor.

referred to as the *block* Gram-Schmidt preprocessor since it performs the orthogonalization of the estimation subspace in blocks.

Efficient simultaneous implementation of multiple beamformers is possible with the modular structure. To illustrate, let  $C_1$  and  $C_2$  denote constraint matrices of dimension  $N \times L_1$  and  $N \times L_2$  respectively for beamformer1 and beamformer2 with  $f_1$  and  $f_2$  the corresponding response vectors. Define  $U = [C_1, C_2]$  and  $L = L_1 + L_2$ . Assuming  $L < N$ , the orthogonal complement of the subspace spanned by the columns of  $U$  is spanned by the columns of an  $N \times (N - L)$  matrix  $\bar{U}$ . It follows that the orthogonal complement of the span of the columns of  $C_1$  is spanned by the columns of the matrix  $C_{n_1} = [\bar{U}, D_1]$ , and the orthogonal complement of the span of the columns of  $C_2$  is spanned by the columns of the matrix  $C_{n_2} = [\bar{U}, D_2]$  for appropriately chosen  $D_1$  and  $D_2$ . These two beamformers are implemented simultaneously using the modular decomposition structure shown in Fig. 6. In this figure,  $w_{q1}$  and  $w_{q2}$  are  $N \times 1$  column vectors that satisfy the constraint equations for the two beamformers. It is straightforward to extend this approach and simultaneously implement more than two beamformers.

Simultaneous implementation of multiple beamformers is more computationally efficient than separate implementation because the adaptive degrees of freedom common to each beamformer are updated by the common preprocessor. For instance, if the two beamformers of the previous example are implemented separately using the modular decomposition, a total of  $N(2N - L)$  adaptive weights need to be computed or updated at each iteration of an adaptive algorithm. In the implementation described in Fig. 6, there are a total of  $N^2$  adaptive weights that need to be updated or computed at each iteration of an adaptive algorithm. Thus, this implementation offers a savings of  $N(N - L)$  adaptive weights.

Note that simultaneous implementation of multiple beamformers is also possible with the Gram-Schmidt preprocessor. The block diagram shown in Fig. 7 depicts a Gram-Schmidt preprocessor that simultaneously implements the two beamformers described previously. In Fig. 7, the common preprocessor consists of  $N - L$  stages that are identical in form to the initial stages of the Gram-Schmidt preprocessor of Fig. 4. Let  $u_c = \bar{U}^H x$ ,  $\mathcal{M}_c = \text{span}\{u_c\}$  and  $u_{12} =$

$[D_1, w_{q1}, D_2, w_{q2}]^H x$ , then the output of the common preprocessor is a dimension  $(L + 2) \times 1$  random vector  $x_c = \mathcal{P}_\perp(u_{12} | \mathcal{M}_c)$ . The first  $L_2 + 1$  elements of this vector are inputs to preprocessor1 and the remaining  $L_1 + 1$  are inputs to preprocessor2. There are  $L_2$  stages in preprocessor1 and  $L_1$  stages in preprocessor2; both are identical in form to the initial stages of the Gram-Schmidt preprocessor. If  $[u_1, d_1, u_2, d_2]^T = x_c$  where  $u_1$  is  $L_2 \times 1$  and  $u_2$  is  $L_1 \times 1$ , then  $y_1 = \mathcal{P}_\perp(d_1 | \text{span}\{u_1\})$  and  $y_2 = \mathcal{P}_\perp(d_2 | \text{span}\{u_2\})$  are the outputs of preprocessor1 and preprocessor2, that is  $y_1$  and  $y_2$  are the outputs of beamformer1 and beamformer2 respectively. The simultaneous Gram-Schmidt preprocessor implementation of Fig. 7 uses  $(N - L)(N - L - 1)/2$  fewer adaptive weights than separate Gram-Schmidt preprocessor implementation.

REFERENCES

- [1] T. C. Liu and B. D. Van Veen, "A modular structure for the implementation of linearly constrained minimum variance beamformers," *IEEE Trans. Signal Processing*, vol. 39, no. 10, pp. 2343-2346, Oct. 1991.
- [2] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.
- [3] R. A. Mozingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
- [4] B. D. Van Veen, "Minimum variance beamforming," in *Adaptive Radar Detection*, S. Haykin and A. Stienhardt, Eds. New York: Wiley, 1992.
- [5] G. H. Golub and C. V. Van Loan, *Matrix Computation*. Baltimore, MD: John Hopkins Press, 1988.

On Constructing Regular Filter Banks from Domain Bounded Polynomials

L. M. G. M. Tolhuizen, I. A. Shah, and A. A. C. M. Kalker

**Abstract**—The design of regular two-channel bi-orthogonal filter banks is shown to be reducible to the design of pairs of real polynomials with domain bounded to the interval [-1,1]. Techniques for designing polynomials satisfying various constraints are outlined. Transformation of polynomials to multidimensional bi-orthogonal filter banks is presented.

I. INTRODUCTION

In this correspondence we will be concerned with the design of critically sampled, 2-channel, zero phase bi-orthogonal (BO) filter banks (FB) [3], [17]. The approach we follow is to reduce the design of 1-D zero-phase BO FB to the design of appropriate real polynomials on the interval [-1,1]. These polynomials are then transformed to yield 1-D and also multidimensional (m-D) zero-phase BO FB. A 1-D BO filter bank of the above class is characterized by the pair of filters  $\{G_0, H_0\}$  in the low-pass channel. The filters in the high-pass channel are up to an even delay given by  $\{z^{-1}G_1(z), zH_1(z)\} = \{z^{-1}H_0(-z), zG_0(-z)\}$ . With this restriction, the perfect reconstruction (PR) property of the filter bank

Manuscript received December 28, 1992; revised July 5, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. Sergio D. Cabrera.

The authors are with Philips Research, 5656 AA, Eindhoven, The Netherlands.

IEEE Log Number 9214189.