638–640, May 1990.
[18] G. Seiler and J. A. Nossek, "Winner-take-all cellular neural networks," *IEEE Trans. Circuits Syst. II,* vol. 40, pp. 184–190, Mar. 1993.
[19] G. Seiler, A. J. Schuler, and J. A. Nossek, "Design of robust cellular neural networks," *IEEE Trans. Circuits Syst. I,* vol. 40, pp. 358–364, May 1993.
[20] T. Szirányi and J. Csicsvári, "High-speed character recognition using a dual cellular neural network architecture (CNND)," *IEEE Trans. Circuits Syst. II,* vol. 40, pp. 223–231, Mar. 1993.

Weighted Low-Rank Approximation of General Complex Matrices and Its Application in the Design of 2-D Digital Filters

W.-S. Lu, S.-C. Pei, and P.-H. Wang

***Abstract*—In this brief we present a method for the weighted low-rank approximation of general complex matrices along with an algorithmic development for its computation. The method developed can be viewed as an extension of the conventional singular value decomposition to include a nontrivial weighting matrix in the approximation error measure. It is shown that the optimal rank-*K* weighted approximation can be achieved by computing *K* generalized Schmidt pairs and an iterative algorithm is presented to compute them. Application of the proposed algorithm to the design of FIR two-dimensional (2-D) digital filters is described to demonstrate the usefulness of the algorithm proposed.**

***Index Terms*—2-D digital filters, singular value decomposition.**

I. Introduction

As one of the basic and important tools in numerical linear algebra, the singular value decomposition (SVD) [1]–[3] has found numerous scientific and engineering applications in the past. An excellent outline on its applications in linear algebra and linear systems can be found in [4]. Sample applications of the SVD in automatic control, robotics, image processing, reduced-rank signal processing, and design of two-dimensional (2-D) digital filters can also be found in [5]–[16]. In a filter design context, the SVD method [10]–[16] starts with a complex matrix F obtained by sampling the desired frequency response, and the application of SVD to F allows one to decompose a complex 2-D design task into a set of simple 1-D design tasks with guaranteed design accuracy. An important property of the SVD utilized in this regard is that the SVD of F of rank r offers a series of optimal low-rank approximations of F in both Euclidean and Frobenius norm sense. That is, if

$$F = U\Sigma V^H = \sum_{i=1}^{r} \sigma_i u_i v_i^H \tag{1}$$

is a SVD of F, then for *any* K between 1 and r,

$$\min_{\operatorname{rank}(\hat{F}_K)=K} \|F - \hat{F}_K\|_{2,F} = \|F - F_K\|_{2,F} \tag{2}$$

where

$$F_K = \sum_{i=1}^{K} \sigma_i u_i v_i^H. \tag{3}$$

Although the SVD method has become a successful design tool, a weak point of the method is that it treats all entries of the sampled frequency response matrix *equally,* which could in some cases lead to degraded designs. In order to discriminate between the important and unimportant portions of the matrix, we seek to find a low-rank approximation of F such that for a fixed K with $1 < K < r$, the rank K matrix

$$F_K = \sum_{i=1}^{K} \sigma_i u_i v_i^H \tag{4}$$

best approximates F in the *weighted* Frobenius norm sense. That is

$$\min_{\operatorname{rank}(\hat{F}_K)=K} \|W \circ (F - \hat{F}_K)\|_F = \|W \circ (F - F_K)\|_F \tag{5}$$

where W is a weighting matrix with the same size as F, $W \circ Y$ denotes the entrywise multiplication of W with Y, which is often termed as Hadamard or Schur product in the literature. In the rest of the brief, we shall call (4), (5) a weighted rank K approximation of F.

In the literature the weighted low rank approximation (WLRA) problem was considered by Shpak [16] in a filter design context for a real matrix F. His approach is to treat (5) as a numerical minimization problem so that the conventional optimization techniques [17], [18] can be used to find a solution. However, the optimization involved requires a large amount of computation, particularly when u_i and v_i are of high dimension. The objectives of this brief are twofold. First, we investigate in Section II the WLRA for a general complex matrix $F \in C^{m\times n}$. It is shown that for a fixed K (which is the rank of F_K approximating F), the WLRA can be characterized by K generalized Schmidt pairs which are nonlinear extension of the conventional Schmidt pairs obtained by the SVD of F. We present an iterative algorithm for numerical computation of the generalized Schmidt pairs. Convergence and computation complexity issues of the algorithm are addressed. Also proposed in Section II is a suboptimal solution to the WLRA problem. This suboptimal WLRA (S-WLRA) is obtained by computing one pair of vectors u_i and v_i at a time, leading to considerably reduced computation complexity and hence offers a feasible solution to those approximation problems where the matrix F is of high dimension. As the second objective of the brief, the S-WLRA is applied to design FIR 2-D digital filters. In Section III, the S-WLRA is applied to design linear phase FIR 2-D filters. An example is included to illustrate the design algorithm and to compare the WLRA-SVD method with the conventional SVD method.

II. Weighted Low-Rank Approximation of Complex Matrices

A. Preliminaries

The singular value decomposition of a rectangular complex matrix $F \in C^{m\times n}$ is the decomposition (1) where $U \in C^{m\times m}$, $V \in C^{n\times n}$ are unitary and

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$$

where $\Sigma_r = \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\}$ with singular values $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$, $r = \operatorname{rank}(F)$. Writing $U = [u_1 \cdots u_r \cdots u_m]$ and

Manuscript received April 25, 1995; revised September 10, 1996. This work was supported by the National Science Council under Grant NSC84-2213-E002-076. This paper was recommended by Associate Editor H. Reddy.

W.-S. Lu is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, B.C., Canada V8W 3P6.

S.-C. Pei and P.-H. Wang are with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

Publisher Item Identifier S 1057-7122(97)03515-0.

$V = [v_1 \cdots v_r \cdots v_n]$, (1) implies that for $1 \le i \le r$

$$F v_i = \sigma_i u_i \tag{6}$$

$$F^H u_i = \sigma_i v_i \tag{7}$$

with

$$\|u_i\|_2 = \|v_i\|_2 = 1 \tag{8}$$

In the literature, the first pair of vectors $\{u_1, v_1\}$ associated with the largest singular value σ_1 is called the Schmidt pair of F [20], [21]. For the sake of convenience we, in the rest of the brief, shall call $\{u_i, v_i\}$ (for $1 \le i \le r$) the ith Schmidt pair of F. From (1) we see that the r Schmidt pairs along with the associated singular values are sufficient to characterize matrix F

$$F = \sum_{i=1}^{r} \sigma_i u_i v_i^H. \tag{9}$$

From a filter-design point of view, the following form of the SVD is often used

$$F = \sum_{i=1}^{r} \tilde{u}_i \tilde{v}_i^H \tag{10}$$

where $\tilde{u}_i = \sigma_i^{\frac{1}{2}} u_i$ and $\tilde{v}_i = \sigma_i^{\frac{1}{2}} v_i$ can be interpreted as 1-D frequency responses when F is a sampled 2-D frequency response [14], [15].

An important property of SVD is the Eckart-Young theorem [2] described by (2) and (3). This means that if K pairs of 1-D transfer functions $f_i(z_1)$ and $g_i(z_2)$ are found such that the sampled frequency response matrix from $f_i(z_1)g_i(z_2)$ perfectly matches $\tilde{u}_i \tilde{v}_i^H$, then $\sum_{i=1}^{K} f_i(z_1)g_i(z_2)$ would be an optimal design in the Euclidean or Frobenius norm sense when the number of parallel sections used is limited to K. Another useful property of the SVD is that the SVD of matrix $F - u_1 v_1^H$ is

$$F - u_1 v_1^H = U_1 \Sigma_1 V_1^H$$

where

$$U_1 = [u_2 \cdots u_r \cdots] \quad \text{and} \quad V_1 = [v_2 \cdots v_r \cdots]$$

are column-orthogonal, and

$$\Sigma_1 = \begin{bmatrix} \Sigma_{2r} & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_{2r} = \operatorname{diag}\{\sigma_2, \ldots, \sigma_r\}.$$

Consequently, the Eckart–Young theorem implies that

$$\min_{\operatorname{rank}(\hat{F}_K)=K} \left\| \left(F - u_1 v_1^H\right) - \hat{F}_K \right\|_{2,F} = \left\| F - F_K^{(1)} \right\|_{2,F} \tag{11}$$

where

$$F_K^{(1)} = \sum_{i=2}^{K+1} \sigma_i u_i v_i^H = \sum_{i=2}^{K+1} \tilde{u}_i \tilde{v}_i^H \tag{12}$$

for $1 \le K \le r-1$. It follows from (2), (3) and (11), (12) that the pth Schmidt pair of F can be obtained as the *first* Schmidt pair of $F - \sum_{i=1}^{p-1} \sigma_i u_i v_i^H$.

B. Weighted Low-Rank Approximation of a Complex Matrix

In what follows we show that the optimal solution to the weighted rank-K approximation of a complex matrix F, which is defined by (4) and (5), is characterized by K generalized Schmidt pairs that can be viewed as a nonlinear version of (6) and (7). For a given complex matrix $F \in C^{m \times n}$ and a real weighting $W \in R^{m \times n}$, we consider the error function for the rank-K approximation problem given by

$$J = \left\| W \circ \left(F - \sum_{i=1}^{K} u_i v_i^H \right) \right\|_F^2. \tag{13}$$

Denoting

$$F_W = W \circ F$$

and

$$R_i = W \circ \left(u_i v_i^H\right)$$

(13) can be written as

$$J = \operatorname{tr}\left[\left(\sum R_i^H\right)\left(\sum R_i\right)\right] - \operatorname{tr}\left(F_W^H \sum R_i\right) - \operatorname{tr}\left[\left(\sum R_i^H\right) F_W\right] + c$$

where $\operatorname{tr}[\cdot]$ denotes the trace of the matrix involved, and c is a constant. Computing the gradient of J with respect to these variables is straightforward but tedious, and is omitted here. By letting the gradient be zero, we obtain the following system of nonlinear equation (14) and (15) as shown at the bottom of the page, where u and v are the mK- and nK-dimensional column vectors defined by

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_K \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_K \end{bmatrix} \tag{16}$$

respectively, P_u and P_v are the permutation matrices of size $mK \times mK$ and $nK \times nK$ such that

$$P_u u = \begin{bmatrix} u_{11} \\ \vdots \\ u_{K1} \\ \vdots \\ u_{1m} \\ \vdots \\ u_{Km} \end{bmatrix}, \quad P_v v = \begin{bmatrix} v_{11} \\ \vdots \\ v_{K1} \\ \vdots \\ v_{1n} \\ \vdots \\ v_{Kn} \end{bmatrix}$$

with u_{ij} and v_{ij} representing the jth entry of u_i and v_i, respectively, $a_{ki}^{(r_l)}$ and $b_{ki}^{(c_l)}$ are the inner products defined by

$$a_{ki}^{(r_l)} = \left\langle w_{rl}^T \circ v_k,\ w_{rl}^T \circ v_i \right\rangle \tag{17}$$

and

$$b_{ki}^{(c_l)} = \left\langle w_{cl} \circ u_k,\ w_{cl} \circ u_i \right\rangle \tag{18}$$

$$(W \circ F_W) v_i = \begin{bmatrix} a_{1i}^{(r_1)} & \cdots & a_{Ki}^{(r_1)} & 0 & \cdots & \cdot & \cdot & \cdots & 0 \\ 0 & \cdots & 0 & a_{1i}^{(r_2)} & \cdots & a_{Ki}^{(r_2)} & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & \cdot & \cdot & \cdots & \cdot & a_{1i}^{(r_m)} & \cdots & a_{Ki}^{(r_m)} \end{bmatrix} P_u u \triangleq \psi_{r_i} u \tag{14}$$

$$(W \circ F_W)^H u_i = \begin{bmatrix} b_{1i}^{(c_1)} & \cdots & b_{Ki}^{(c_1)} & 0 & \cdots & \cdot & \cdot & \cdots & 0 \\ 0 & \cdots & 0 & b_{1i}^{(c_2)} & \cdots & b_{Ki}^{(c_2)} & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & \cdot & \cdot & \cdots & \cdot & b_{1i}^{(c_m)} & \cdots & b_{Ki}^{(c_m)} \end{bmatrix} P_v v \triangleq \psi_{c_i} v \tag{15}$$

with w_{rl} and w_{cl} being the lth row and column of W, respectively.

Further, by defining block diagonal matrix

$$\phi = \mathrm{diag}\,\{\overbrace{W \circ F_W, \ldots, W \circ F_W}^{K \text{ blocks}}\} \tag{19}$$

and

$$\psi_r = \begin{bmatrix} \psi_{r1} \\ \vdots \\ \psi_{rK} \end{bmatrix} \quad \text{and} \quad \psi_c = \begin{bmatrix} \psi_{c1} \\ \vdots \\ \psi_{cK} \end{bmatrix} \tag{20}$$

the K sets of (14) and (15) with $i = 1 \ldots, K$ can be put together as

$$\phi v = \psi_r u \tag{21}$$

$$\phi^H u = \psi_c v \tag{22}$$

Equations (21) and (22) are important as they characterize the optimal u and v, and therefore the optimal u_i and v_i (for $i = 1, \ldots, K$) via (16) that minimize the error function J. As can be seen from (19), matrix ϕ is independent of parameters u and v. However, the nonzero entries in ψ_r and ψ_c are in general dependent on u_i and v_i quadratically as is evidenced from (14), (15), (17), and (18). In spite of the nonlinear nature of (21), (22), it is worthwhile to notice the analogy between (21), (22), and (6), (7). As a matter of fact, if W is a trivial weighting with all entries being the same constant, it can readily be verified that the $\{\tilde{u}_i, \tilde{v}_i, i = 1, \ldots, K\}$ obtained from the SVD of F offers a solution to (21) and (22). It is for this reason that we shall in the sequel call $\{u_i/\|u_i\|,\ v_i/\|v_i\|,\ i = 1, \ldots, K\}$ determined by (21), (22), and (16) the ith *generalized Schmidt pair* for the given $F,\ W$, and K.

C. Computation of Generalized Schmidt Pairs

A Recursive Computation Scheme and a Convergence Analysis: To obtain the K generalized Schmidt pairs, the $(m+n)K$ nonlinear equations defined by (21) and (22) need to be solved. In what follows we propose a scheme for recursively computing u and v. It starts by writing (21) and (22) as

$$u = \psi_r^{-1}(u, v)\phi v \tag{23a}$$

$$v = \psi_c^{-1}(u, v)\phi^H u \tag{23b}$$

respectively, where the dependence of ψ_r and ψ_c on u and v is explicitly indicated. With an initial $u^{(0)},\ v^{(0)}$, (23) suggests a scheme to compute

$$p^{(k)} = \psi_r^{-1}(u^{(k)}, v^{(k)})\phi v^{(k)} \tag{24a}$$

$$q^{(k)} = \psi_c^{-1}(u^{(k)}, v^{(k)})\phi^H u^{(k)} \tag{24b}$$

and then to obtain $u^{(k+1)},\ v^{(k+1)}$ through a linear combination of $u^{(k)},\ v^{(k)}$ with $p^{(k)},\ q^{(k)}$ as

$$u^{(k+1)} = \alpha u^{(k)} + (1-\alpha)p^{(k)} \tag{25a}$$

$$v^{(k+1)} = \alpha v^{(k)} + (1-\alpha)q^{(k)} \tag{25b}$$

where $\alpha \in (0, 1)$ is a relaxation parameter. Denote

$$x^{(k)} = \begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix}. \tag{26}$$

Equation (24) can be expressed as

$$\begin{bmatrix} p^{(k)} \\ q^{(k)} \end{bmatrix} = \begin{bmatrix} 0 & \psi_r^{-1}(u^{(k)}, v^{(k)})\phi \\ \psi_c^{-1}(u^{(k)}, v^{(k)})\phi^H & 0 \end{bmatrix} \begin{bmatrix} u^{(k)} \\ v^{(k)} \end{bmatrix} \equiv \Phi(x^{(k)}). \tag{27}$$

Combining (25) with (26) gives

$$x^{(k+1)} = \Psi(x^{(k)}) \tag{28a}$$

with

$$\Psi(x) = \alpha x + (1-\alpha)\Phi(x). \tag{28b}$$

By (28a) we see that if the sequence $\{x^{(k)}, k = 0, 1, \ldots\}$ converges to vector, say x^*, then x^* is a *fixed point* of function $\Psi(x)$, i.e., $\Psi(x^*) = x^*$. By (28b), this fixed point satisfies

$$x^* = \alpha x^* + (1-\alpha)\Phi(x^*)$$

which gives

$$x^* = \Phi(x^*)$$

From (27) it follows that $x^* = [u^{*T}\ v^{*T}]^T$ satisfies (23) and therefore contains K generalized Schmidt pairs.

A sufficient condition for the convergence of $\{x^{(k)}\}$ generated by (28) is the existence of a constant $\beta \in (0, 1)$ and a positive integer L such that

$$\|\Phi(x^{(k)}) - \Phi(x^{(k-1)})\| \le \beta \|x^{(k)} - x^{(k-1)}\| \tag{29}$$

for $k \ge L$, where $\|\cdot\|$ denotes the Euclidean norm. To show this, note that by (28b) and (29) we have for $k \ge L$

$$\|\Psi(x^{(k)}) - \Psi(x^{(k-1)})\| \le \gamma \|x^{(k)} - x^{(k-1)}\| \tag{30}$$

where $\gamma = \alpha + (1-\alpha)\beta < 1$. In conjunction with (28a), (30) implies that

$$\|\Psi(x^{(k)}) - \Psi(x^{(k-1)})\| \le \gamma^{k-L+1}\|x^{(L)} - x^{(L-1)}\|.$$

So for sufficiently large integers m and n with $m > n \ge L$, we have

$$\begin{aligned} \|x^{(m)} - x^{(n)}\| &\le \sum_{l=1}^{m-n-1} \|x^{(n+l+1)} - x^{(n+l)}\| \\ &\le \sum_{l=0}^{m-n-1} \gamma^{n+l-L+1}\|x^{(L)} - x^{(L-1)}\| \\ &= \frac{\gamma^{n-L+1} - \gamma^{m-L+1}}{1-\gamma}\|x^{(L)} - x^{(L-1)}\| \end{aligned}$$

which approaches to zero when $m, n \to \infty$. Therefore $\{x^{(k)}\}$ is a Cauchy sequence in a finite-dimensional Euclidean space, whose convergence is guaranteed.

Furthermore, by (28a) we see that (30) is equivalent to

$$\|x^{(k+1)} - x^{(k)}\| \le \gamma \|x^{(k)} - x^{(k-1)}\|$$

i.e.,

$$\eta_k \le \gamma \tag{31}$$

for $k > L$, where η_k is the ratio $\eta_k = \|x^{(k+1)} - x^{(k)}\|/\|x^{(k)} - x^{(k-1)}\|$. In words, we conclude that the sequence $\{x^{(k)}\}$ generated from (24) and (25) converges if the ratio η_k has a less-than-unity upper bound for $k \ge L$ where L is a positive integer.

Functions Φ and Ψ satisfying (29) and (30) with $\beta < 1$ and $\gamma < 1$ are called *contraction mappings*. With this term the above analysis may be summarized as three sufficient conditions, each of which ensures the convergence of sequence $\{x^{(k)}\}$ generated recursively by (24), (25):

1) $\Phi(x)$ is a contraction mapping;
2) $\Psi(x)$ is a contraction mapping;
3) η_k has a less-than-unity upper bound for $k \ge L$.

As a part of the convergence analysis the recursive scheme (24), (25) was applied to a large number of randomly generated matrices F of various sizes $(1 \leq m, n \leq 40)$ along with randomly selected weighting matrices W whose entries are uniformly distributed on [0, 1]. The rank parameter K in the test varies from 1 to $\min(n, m)$. In every case of the test, the convergence of the recursive computation scheme (24), (25) was confirmed by verifying sufficient condition (31). Although further investigation on the specific structure of Ψ and Φ as related to these sufficient conditions remains an interesting issue, this numerical test offers the confidence to use it as a feasible means to compute generalized Schmidt pairs.

We now summarize the proposed recursive computation scheme as in Algorithm 1.

Algorithm 1

Step 1} Select initial $u^{(0)}$ and $v^{(0)}$ and set $k = 0$.
Step 2} Compute $p^{(k)}$ and $q^{(k)}$ using (24).
Step 3} Compute $u^{(k+1)}$ and $v^{(k+1)}$ using (25).
Step 4} Define $x^{(k+1)}$ by (26). If $\|x^{(k+1)} - x^{(k)}\|$ is less than a prescribed tolerance, output $u = u^{(k+1)}, \ v = v^{(k+1)}$, and stop. Otherwise set $k = k+1$ and repeat from Step 2.

Computation Complexity: The most expensive thing in implementing Algorithm 1 is to compute $p^{(k)}$ and $q^{(k)}$ using (24). This is equivalent to solving the two linear systems of equations

$$\psi_{rk} p = b_v \tag{32a}$$
$$\psi_{ck} q = b_u \tag{32b}$$

for p and q, where $\psi_{rk} = \psi_r(u^{(k)}, v^{(k)})$, $\psi_{ck} = \psi_c(u^{(k)}, v^{(k)})$, $b_v = \phi v^{(k)}$, and $b_u = \phi^H u^{(k)}$. At this point it is important to note that matrices Ψ_r and ψ_c, which are characterized by (14), (15), (17), and (18), are positive definite Hermitian matrices. It is known [3] that such systems can be solved by using stable Cholesky decomposition of the coefficient matrix, and requires $d^3/3$ flops (rather than $2d^3/3$ flops for a general linear system) where d denotes the system's dimension. Hence, solving system (32a) and (32b) requires about $K^3(m^3+n^3)/3$ flops in each iteration when K generalized Schmidt pairs are sought.

The overall computation complexity is therefore $k^* K^3(m^3 + n^3)/3$ flops where k^* is the number of iterations used. There are two factors in the algorithm implementation that are particularly relevant to reducing k^*. The first factor is the selection of the initial vectors $u^{(0)}$ and $v^{(0)}$. Since a generalized Schmidt pair becomes a conventional Schmidt pair when W is a trivial weighting, one may use the conventional Schmidt pair as $u^{(0)}$ and $v^{(0)}$ for the nontrivial weighting case. This choice of $u^{(0)}$ and $v^{(0)}$ considerably reduces the value of k^* compared to a randomly chosen $u^{(0)}$ and $v^{(0)}$ especially when the variations in the entries of W from a trivial one is not large. The second factor is the choice of the relaxation parameter α. It is observed from (25) that the updated $u^{(k+1)}$ and $v^{(k+1)}$ contains $100\alpha\%$ "old" $u^{(k)}$ and $v^{(k)}$. Thus a smaller α would in general lead to a larger k^*. This observation was confirmed in our numerical evaluation or the proposed algorithm, where an α between 0.5 and 0.9 is often found suitable.

D. A Suboptimal Solution of the WLRA Problem

Unlike the SVD for which the pth Schmidt pair of F is equal to the first Schmidt pair of $F - \sum_{i=1}^{p-1} \sigma_i u_i v_i^H$, the generalized Schmidt pairs defined by (21) and (22) do not in general possess this property unless the weighting W is trivial. Consequently, for a fixed rank K, the K generalized Schmidt pairs cannot be found by recursively minimizing the error function

$$J_{p-1} = \left\| W \circ \left(F_{p-1} - u_p v_p^H\right)\right\|_F \tag{33}$$

with respect to u_p and v_p, for $p = 1, 2, \cdots, K$, where

$$F_{p-1} = F - \sum_{i=1}^{p-1} u_i v_i^H. \tag{34}$$

However, on comparing J_{p-1} in (33) with J in (13) we see that for each p minimizing J_{p-1} involves only $m+n$ complex parameters while minimizing J involves $K(m+n)$ parameters. Evidently, solving K minimization problems (33) and (34) is numerically more feasible especially when F is of high dimension, which is often the case in filter design applications. Of course the resulting rank-K approximation so obtained offers only a suboptimal solution to the WLRA problem.

Quasi-Schmidt Pairs: The u_p and v_p that minimize (33) can be obtained by considering a special case of minimizing J in (13) with $K = 1$ and F replaced by F_{p-1} defined by (34). It can readily be shown that the optimal u_p and v_p satisfy

$$(W \circ F_{p-1}, w) v_p = \psi_r^{(p)} u_p \tag{35}$$
$$(W \circ F_{p-1}, w)^H u_u = \psi_c^{(p)} v_p \tag{36}$$

where

$$F_{p-1,W} = W \circ F_{p-1} \tag{37}$$

$$\psi_r^{(p)} = \begin{bmatrix} \|W_{r1}^T \circ v_p\|^2 & & 0 \\ & \ddots & \\ 0 & & \|W_{rm}^T \circ u_p\|^2 \end{bmatrix} \tag{38}$$

$$\psi_c^{(p)} = \begin{bmatrix} \|W_{c1}^T \circ u_p\|^2 & & 0 \\ & \ddots & \\ 0 & & \|W_{cn}^T \circ u_p\|^2 \end{bmatrix}. \tag{39}$$

The analogy between (35), (36) and (21), (22), and the analogy between (35), (36) and (6), (7) are evident. In fact, with a trivial weighting W (say all its entries are equal to c), we have $W \circ F_{p-1,W} = c^2 F_{p-1}$, $\psi_r^{(p)} = c^2 \|v_p\|^2 I$, and $\psi_c^{(p)} = c^2 \|u_p\|^2 I$. Hence (35) and (36) become

$$F_{p-1} v_p = \|v_p\|^2 u_p$$
$$F_{p-1}^H u_p = \|u_p\|^2 v_p$$

which means that $u_p/\|u_p\|$ and $v_p/\|v_p\|$ is the first Schmidt pair of F_{p-1}, therefore the pth Schmidt pair of F! It is for this reason we shall call $u_p/\|u_p\|$, $v_p/\|v_p\|$, characterized by (35), (36) *the* kth *quasi-Schmidt pair*. To compute the quasi-Schmidt pairs, Algorithm 1 is applicable with (24a) and (24b) replaced by

$$p^{(k)} = \left[\psi_r^{(p)}\left(W, v_p^{(k)}\right)\right]^{-1} \left(W \circ F_{p-1,W}\right) v_p^{(k)} \tag{40a}$$
$$q^{(k)} = \left[\psi_c^{(p)}\left(W, u_p^{(k)}\right)\right]^{-1} (W \circ F_{p-1,W})^H u_p^{(k)} \tag{40b}$$

and (25a) and (25b) replaced by

$$u_p^{(k+1)} = \alpha u_p^{(k)} + (1-\alpha) p^{(k)} \tag{41a}$$
$$v_p^{(k+1)} = \alpha v_p^{(k)} + (1-\alpha) q^{(k)} \tag{41b}$$

respectively.

Generalized versus Quasi-Schmidt Pairs—Some Comparisons: Since both $\psi_r^{(p)}$ and $\psi_c^{(p)}$ are diagonal matrices, it can readily be verified that the overall computation complexity for evaluating K quasi-Schmidt pairs using the recursive scheme (40), (41) is $\sum_{i=1}^{K} k_i^*(3mn + m + n)$ flops where k_i^* denotes the number of iterations used to compute the ith quasi-Schmidt pair. Let $k^* = \max\{k_i^*, \ i = 1, \cdots, K\}$, then the overall computation complexity is upper-bounded by $k^* K(3mn + m + n)$ flops which is substantially less than that of computing K generalized Schmidt pair (requiring $k^* K^3(m^3+n^3)/3$ flops) even for moderate $m, \ n$, and K.

The improved computation efficiency is, however, at the expense of quality degradation. In what follows we give a qualitative account of how the degradation is related to the weighting W and parameter K. If we denote the K quasi-Schmidt pairs by $\hat{u}_i,\ \hat{v}_i,\ i = 1,\cdots,K$, then

$$\begin{aligned}
&\min_{\substack{u_i,v_i\\ i=1,\ldots,K}} \left\| W \circ \left(F - \sum_{i=1}^{K} u_i v_i^H \right) \right\| \\
&\le \min_{\substack{u_i,v_i\\ i=2,\ldots,K}} \left\| W \circ \left[\left(F - \hat{u}_1 \hat{v}_1^H \right) - \sum_{i=2}^{K} u_i v_i^H \right] \right\| \\
&\le \min_{\substack{u_i,v_i\\ i=3,\ldots,K}} \left\| W \circ \left[\left(F - \hat{u}_1 \hat{v}_1^H - \hat{u}_2 \hat{v}_2^H \right) - \sum_{i=3}^{K} u_i v_i^H ht \right] \right\| \\
&\le \cdots \\
&\le \min_{u_K,v_K} \left\| W \circ \left[\left(F - \sum_{i=1}^{K-1} \hat{u}_i \hat{v}_i^H \right) - u_K v_K^H \right] \right\| \\
&= \left\| W \circ \left(F - \sum_{i=1}^{K} \hat{u}_i \hat{v}_i^H \right) \right\|. \qquad (42)
\end{aligned}$$

The minimum of the left-end term in (42) is achieved by the K generalized Schmidt pairs while the minimum of the term on the right-hand side of the last inequality in (42) is achieved by the K quasi-Schmidt pairs, and there are $K-1$ inequalities in between. As was noted in Sections II-B and II-D-1, with a trivial W the generalized Schmidt pairs and quasi-Schmidt pairs for any given K are identical and equal to the conventional Schmidt pairs. In such a case, all inequalities in (42) becomes equality. Stating it in another way, the inequalities would hold *strictly* if W is nontrivial and F is nonsparse in the sense that $F \circ W$ differs from cF for any constant C. Under these circumstances (42) clearly indicates that the quality degradation of the suboptimal solution deepens with parameter K and the nontriviality of weighting W.

III. Design of Linear Phase FIR 2-D Filters Using WLRA

A. The Method

Let F be the sampled frequency response matrix obtained by sampling the desired frequency response. If linear phase response is required in the design, one only needs to sample the desired amplitude response of the filter since the linearity of the phase response is guaranteed if each 1-D FIR filter involved is of linear phase (this will become more apparent shortly). In this case F is a real-valued matrix. Furthermore, if the desired filter is quadrantally symmetric, a quarter (say the lower-right block) of F is sufficient to characterize the whole matrix. So for a typical sampling density of 61 × 61 over the baseband, one can work on the WLRA problem for a real matrix of 31 × 31 if the filter to be designed is quadrantally symmetric. Denoting this portion of F and W by F and W again, the next step is to determine an adequate value of K. Without loss of generality we assume that the entries of W fall over interval [0, 1]. It follows that for any matrix $\hat{F}_K$ of rank K.

$$\|W \circ (F - \hat{F}_K)\|_F \le \|F - \hat{F}_K\|_F$$

which leads to an upper bound for the WLRA problem as

$$\min_{\operatorname{rank}(\hat{F}_K)=K} \|W \circ (F - \hat{F}_K)\|_F \le \min_{\hat{F}_K} \|F - \hat{F}_K\|_F = \left(\sum_{i=K+1}^{r} \sigma_i^2 \right)^{\frac{1}{2}} \qquad (43)$$

where $\{\sigma_i : i = K+1,\cdots,r\}$ are the last $r-K$ singular values of F. Obviously, this upper bound is quite tight if W does not severely deviate from a trivial weighting. One can then choose a K such that $(\sum_{i=K+1}^{r} \sigma_i^2)^{\frac{1}{2}}$ is small. For the cases where W is far from trivial, the K determined above can serve as a preliminary choice of K and a smaller K might be reached by a trial-and error approach.

Having determined the value of K, Algorithm 1 can be used to find the K generalized or quasi-Schmidt pairs. For an F of dimension 31 × 31, typical number of nonzero singular values are in the range of 10 to 20. If $K = 10$ is chosen, the evaluation of the K generalized Schmidt pairs involves inversion of matrices whose size will be as large as $K(m+n) = 620$. On the other hand, the determination of the K quasi-Schmidt pairs can be accomplished by K subproblems each of which has only $m + n = 62$ parameters to work with. In any event, once the K generalized or quasi-Schmidt pairs are computed, the rank K approximation of F is given by

$$F_K = \sum_{i=1}^{K} \sigma_i u_i v_i^H \qquad (44)$$

where the vectors $\sigma_i^{\frac{1}{2}} u_i$ and $\sigma_i^{\frac{1}{2}} v_i$ are used as the desired 1-D frequency responses, and linear phase, FIR and transfer functions $f_i(z_1)$ and $g_i(z_2)$ can be designed whose frequency responses approximate $\sigma_i^{\frac{1}{2}} u_i$ and $\sigma_i^{\frac{1}{2}} v_i$, respectively. The linear-phase, FIR 2-D transfer function is then given by

$$H(z_1, z_2) = \sum_{i=1}^{K} f_i(z_1) g_i(z_2). \qquad (45)$$

B. A Design Example

In this section, the usefulness of the WLRA will be illustrated by designing a circularly symmetric, linear phase, lowpass, FIR 2-D filter. The normalized passband and stopband of the filter are $\omega_p = 0.25$ and $\omega_s = 0.35$. The desired amplitude response is sampled in a density of 61 × 61 over the baseband, and a quarter of the sampled amplitude response matrix is used in the design. The design specifications are that the maximum ripples of the filter in both passband and stopband are less than 0.03. If we use the conventional SVD method [13], then the maximum ripples of a (29, 29) FIR filter are 0.0418 and 0.0181, respectively. The design accuracy in the passband does not meet the requirement while the filter shows a better-than-enough design accuracy in the stopband. Under these circumstances one might expect to be able to design a (29, 29) FIR filter with an adequate weighting W to meet the design requirement.

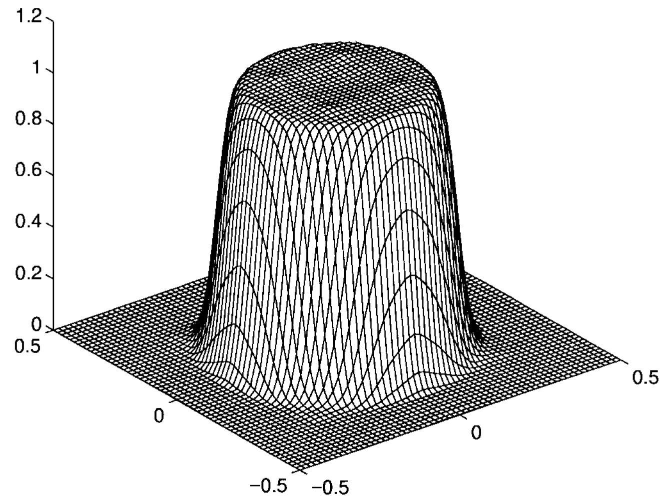

Fig. 1. Amplitude response of the (29, 29) FIR filter.

The weighting matrix used here is a sampled version of the function

$$w(\omega_1, \omega_2) = \begin{cases} 1.2, & 0.225 \le \sqrt{w_1^2 + w_2^2} \le 0.285 \\ 1, & \text{otherwise} \end{cases}$$

where the heavier weights have been given to the region near the passband edge since it is this region larger design error occurs as we have often seen in the conventional SVD-based designs.

Using the suboptimal WLRA algorithm described in Section II-D, twelve quasi-Schmidt pairs are computed and then used to design 1-D transfer functions $f_i(z_1)$ and $g_i(z_2)$. Since both the sampled amplitude response matrix F and weighting matrix W are symmetric, $f_i(z_1)$ and $g_i(z_2)$ have the same coefficients. With $K = 12$, the 2-D transfer function can be found using (42). The maximum ripples of the filter designed in the passband and stopband are 0.0296 and 0.0294, respectively. The amplitude response of the filter is depicted in Fig. 1.

References

[1] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, U.K.: Cambridge Univ. Press, 1985.
[2] G. W. Stewart, *Introduction to Matrix Computations*. New York: Academic, 1974.
[3] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.
[4] V. C. Klema and A. J. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Trans. Automatic Contr.*, vol. 25, pp. 164–176, Apr. 1980.
[5] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. New York: Macmillan, 1992.
[6] T. Yoshikawa, *Foundations of Robotics*. Cambridge, MA: MIT Press, 1990.
[7] Y. Nakamura, *Advanced Robotics*. Reading, MA: Addison-Wesley, 1991.
[8] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
[9] L. L. Scharf, "The SVD and reduced rank signal processing," *Signal Processing*, vol. 25, pp. 113–133, Nov. 1991.
[10] S. Treitel and J. L. Shanks, "The design of multistage separable planar filters," *IEEE Trans. Geosci. Electron.*, vol. GE-9, pp. 10–27, 1971.
[11] R. E. Twogood and S. K. Mitra, "Computer-aided design of separable two-dimensional digital filters," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 25, pp. 165–169, Feb. 1977.
[12] A. Antoniou and W.-S. Lu, "Design of 2-D digital filters by using the singular value decomposition," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 1191–1198, Oct. 1987.
[13] W.-S. Lu, H.-P. Wang, and A. Antoniou, "Design of 2-D digital filters using the singular value decomposition and balanced approximation," *IEEE Trans. Signal Processing*, vol. 39, pp. 2253–2262, Oct. 1991.
[14] T. B. Deng and M. Kawamata, "Frequency domain design of 2-D digital filters using the iterative singular value decomposition," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1225–1228, Oct. 1991.
[15] N. Yazdi, M. Ahmadi, and J. J. Soltis, "Efficient design of 2-D linear phase FIR filters using singular value decomposition," *Elect. Lett.*, vol. 38, no. 24, pp. 2256–2258, Nov. 1992.
[16] D. Shpak, "A weighted-least-squares matrix decomposition method with application to the design of two-dimensional digital filters," in *Proc. IEEE Midwest Symp. Circuits Syst.*, Calgary, Canada, Aug. 1990, pp. 1070–1073.
[17] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1984.
[18] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. Chichester, U.K.: Wiley, 1987.
[19] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge, U.K.: Cambridge Univ. Press, 1991.
[20] V. M. Adamjan, D. Z. Arov, and M. G. Krein, "Analytic properties of Schmidt pairs for Hankel operator and the generalized Schur–Takagi problem," *Math. USSR Sbornik*, vol. 15, pp. 31–73, 1971.
[21] ——, "Infinite block Hankel matrices and relatd extension problems," *AMS Transl.*, vol. 111, pp. 133–156, 1978.