Hermite-Gaussian Like Eigenvectors of the Discrete Fourier Transform Matrix Based on the Singular Value Decomposition of its Orthogonal Projection Matrices

Magdy Tawfik Hanna¹, Nabila Philip Attalla Seif² and Waleed Abd El Maguid Ahmed³

Abstract

A new technique is proposed for generating initial orthonormal eigenvectors of the Discrete Fourier Transform matrix F by the singular value decomposition of its orthogonal projection matrices on its eigenspaces and efficiently computable expressions for those matrices are derived. In order to generate Hermite-Gaussian like orthonormal eigenvectors of F given the initial ones, a new method called the Sequential Orthogonal Procrustes Algorithm (SOPA) is presented based on the sequential generation of the columns of a unitary matrix rather than the batch evaluation of that matrix as in the Orthogonal Procrustes Algorithm (OPA). It is proved that for any of the SOPA, the OPA, or the Gram-Schmidt Algorithm (GSA) the output Hermite-Gaussian like orthonormal eigenvectors are invariant under the change of the input initial orthonormal eigenvectors.

Index Terms: Discrete fractional Fourier transform, Hermite-Gaussian like eigenvectors, orthogonal procrustes algorithm, Gram-Schmidt algorithm, singular value decomposition.

¹ Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University / Fayoum Branch, Fayoum 63514, Egypt, Tel: 0020-2-5706033, Fax: 0020-2-5706033, Email: hanna@ieee.org
² Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University, Giza, Egypt, Email: npaseif@mail.com
³ Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University / Fayoum Branch, Fayoum 63514, Egypt.
I. INTRODUCTION

Having developed the continuous FRactional Fourier Transform (FRFT) [1-3], current research is taking place to develop its discrete counterpart, namely the Discrete FRactional Fourier Transform (DFRFT) [4-8]. In order for the DFRFT to satisfy the requirements of unitarity and index additivity, orthonormal eigenvectors should be generated for the Discrete Fourier Transform (DFT) matrix F. In order for the DFRFT to approximate its continuous counterpart, it is logical to demand that the eigenvectors of F approximate the Hermite-Gaussian functions which are the eigenfunctions of the FRFT [9].

Candan, Kutay and Ozaktas obtained a second order difference equation by discretizing the second order differential equation satisfied by the Hermite-Gaussian functions [6,7]. Periodic solutions of this difference equation exist since its coefficients are periodic with period N. One period of each solution sequence forms the elements of an eigenvector of an almost tridiagonal real symmetric matrix S. The matrix S and the DFT matrix F have a common set of eigenvectors because they commute. Candan et. al. used those Hermite-Gaussian like orthonormal eigenvectors of F as a basis for a legitimate definition of the DFRFT. Actually the work of Candan et. al. was an extension of that of Dickinson and Steiglitz [10] who previously studied the eigenstructure of the special matrix S.

Pei, Yeh and Tseng achieved markedly superior results. They regarded the orthonormal eigenvectors of matrix S only as initial orthonormal basis spanning the eigenspaces of matrix F. In each eigenspace they searched for other orthonormal eigenvectors that better approximate the Hermite-Gaussian functions [5]. The unitarity of matrix F implies that its eigenspaces corresponding to its distinct eigenvalues are orthogonal to each other [11] and the task reduces to finding good Hermite-Gaussian like orthonormal eigenvectors for each eigenspace individually. More specifically since matrix F has four distinct eigenvalues \((- j)^{k-1}, k = 1, \cdots 4\) [12,13], the corresponding initial eigenvectors are grouped as the columns of 4 matrices \(V_k, k = 1, \cdots 4\). Pei et. al. generated a set of n vectors by sampling the Hermite-Gaussian functions and proved that they are approximate eigenvectors of matrix F corresponding to the exact eigenvalues \((- j)^{k-1}, k = 1, \cdots 4\) [4,5]. Those vectors are arranged as the columns of 4 matrices \(U_k, k = 1, \cdots 4\). Pei et. al. proposed two techniques for getting

---

4 Strictly speaking, denoting the matrix \(S\) in the work of Dickinson et. al. [10] and Pei et. al. [5] by \(S_1\) and the matrix \(S\) in the work of Candan et. al. [6] by \(S_2\), the two matrices are related by \(S_2 = S_1 - 4I\). Therefore \(S_1\) and \(S_2\) have the same eigenvectors.
orthonormal eigenvectors of F that better approximate the Hermite-Gaussian functions than the initial orthonormal eigenvectors forming the columns of the matrices $V_k, k = 1, \ldots, 4$ [5].

The first technique is the Gram-Schmidt Algorithm (GSA) where for each value of $k$ separately the columns of $U_k$ are projected on the column space of $V_k$ to get exact nonorthogonal eigenvectors of F that are next orthonormalized by applying the Gram-Schmidt method. The second technique is the Orthogonal Procrustes Algorithm (OPA) where for each value of $k$ separately the desired superior orthonormal eigenvectors are assumed to form the columns of a matrix $\hat{U}_k$ that is expressed as $V_k Q_k$ and the unitary matrix $Q_k$ is evaluated by minimizing the Frobenius norm of the matrix difference $\left(U_k - \hat{U}_k\right)$ [14].

One objective of the current paper is to present a direct technique for generating initial orthonormal eigenvectors of F without appealing to matrix S. The technique depends on the generation of the orthogonal projection matrices of matrix F on its eigenspaces and the computation of initial eigenvectors by applying the singular value decomposition technique. More specifically, expressions are derived for the projections matrices and simplified in order to get efficiently computable forms. A second objective is to present an alternative technique for generating good Hermite-Gaussian like eigenvectors of F given initial ones. In this technique – to be referred to as the Sequential Orthogonal Procrustes Algorithm (SOPA) – the columns of the unitary matrix $Q_k$ are sequentially evaluated by solving a series of constrained minimization problems rather than batch evaluated as in the OPA. A third contribution is the proof that the superior Hermite-Gaussian like eigenvectors computed using any of the three techniques – GSA, OPA or SOPA – are invariant under the change of the initial eigenvectors. This implies that for any of the three refined techniques the final eigenvectors will be the same whether the initial eigenvectors are computed by finding the eigenvectors of the auxiliary matrix S or by the singular value decomposition of the projection matrices. More surprisingly, it will be proved that both the GSA and the SOPA produce identical results despite being algorithmically quite different.

In section II the projection matrices on the four eigenspaces of matrix F will be derived using two completely different methods. In section III each projection matrix is decomposed by the singular value decomposition technique in order to get orthonormal basis of the corresponding eigenspace. In section IV the Gram-Schmidt Algorithm and the Orthogonal Procrustes Algorithm will be surveyed and proved to produce an output that is invariant under the change of the initial eigenvectors. The contributed Sequential Orthogonal
Procrustes Algorithm will be next presented and proved to have the same property. In section V it will be shown that the SOPA and GSA produce identical outputs despite being algorithmically distinct. Some simulation results will be presented in section VI.

II. THE ORTHOGONAL PROJECTION MATRICES ON THE EIGENSPACES OF MATRIX F

The Discrete Fourier Transform matrix $F = (f_{m,n})$ of order $N$ is defined by:

$$f_{m,n} = \frac{1}{\sqrt{N}} W^{(m-1)(n-1)}, \quad m, n = 1, \ldots, N$$

where

$$W = \exp \left( -j \frac{2\pi}{N} \right).$$

Matrix $F$ has the following 4 distinct eigenvalues [12]:

$$\lambda_k = (-j)^{k-1}, \quad k = 1, \ldots, 4.$$

It is straightforward to show that matrix $F$ is unitary and consequently is diagonalizable [15]. According to the spectral theorem [16], $F$ has the following spectral decomposition:

$$F = \sum_{k=1}^{4} \lambda_k P_k$$

where $P_k$ is the orthogonal projection matrix on the $k$th eigenspace of $F$ to be denoted by $E_k$. Two different methods will be presented below for deriving the four projection matrices.

Method A

Since for any integer $m$, matrix $F^m$ has the same eigenvectors and consequently projection matrices as matrix $F$, equation (4) implies that:

$$F^m = \sum_{k=1}^{4} \lambda_k^m P_k, \quad m = 0, 1, \ldots.$$

The special case of $m = 0$ is the resolution of the identity matrix induced by $F$. Since the objective is the derivation of the four projection matrices, the above equation will be written for $m = 0, 1, 2, 3$ in order to get four matrix equations that can be expressed as:
\[
\begin{bmatrix}
P_1 \\
P_2 \\
P_3 \\
P_4 
\end{bmatrix} =
\begin{bmatrix}
I \\
F \\
F^2 \\
F^3 
\end{bmatrix}
\]

where the partitioned matrix of coefficients is given by:

\[
A = \begin{bmatrix}
I & I & I & I \\
\lambda_1 I & \lambda_2 I & \lambda_3 I & \lambda_4 I \\
\lambda_1^2 I & \lambda_2^2 I & \lambda_3^2 I & \lambda_4^2 I \\
\lambda_1^3 I & \lambda_2^3 I & \lambda_3^3 I & \lambda_4^3 I 
\end{bmatrix}
\]

In preparation for solving (6), one obtains the following results from (3):

\[
\sum_{k=1}^{4} \lambda_k^m \lambda_k = \sum_{k=1}^{4} \left( e^{-j0.5\pi \lambda_k} \right)^{k-1} = \sum_{k=1}^{4} \exp\left(-j \frac{2\pi}{4} m(k-1)\right)
\]

\[
= \begin{cases} 
4 & \text{if } m = 4r, \text{ r integer} \\
0 & \text{otherwise}
\end{cases}
\]

(8)

\[
\sum_{k=1}^{4} \lambda_k^m \lambda_k^{m'} = \sum_{k=1}^{4} \exp\left(-j \frac{\pi}{2} (k-1)m\right) \exp\left(j \frac{\pi}{2} (k-1)m\right) = \sum_{k=1}^{4} \exp\left(j \frac{2\pi}{4} (n-m)(k-1)\right)
\]

\[
= \begin{cases} 
4 & \text{if } n-m = 4r, \text{ r integer} \\
0 & \text{otherwise}
\end{cases}
\]

(9)

Using (7)-(9), one obtains:

\[
A A^+ = 4I
\]

(10)

and consequently

\[
A^{-1} = 0.25 A^+. 
\]

(11)

Therefore (6) can be directly solved using (11) and (7) to get the following compact expression for the four orthogonal projection matrices:

\[
P_k = 0.25 \sum_{m=0}^{3} \lambda_k^m F^m, \quad k = 1, \ldots, 4.
\]

(12)

In order to put the above expression in an efficiently computable form, one should utilize the fact that [17, p.351]:

\[
F^2 = \Gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & J \end{pmatrix}
\]

(13)

where J is the contra-identity matrix of order \((N - 1)\) defined by:

5 The superscripts * and + respectively denote the complex conjugate and the complex conjugate transpose.
\[ J = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}. \] (14)

Upon using the fact that [10]:
\[ F^4 = I \] (15)
and the unitarity and symmetry of \( F \), one obtains:
\[ F^3 = F^{-1} = F^* = F^+. \] (16)

From (3), one directly gets:
\[ \lambda_k^2 = (-1)^{(k-1)}. \] (17)

Substituting (13), (16), (3) and (17) in (12), one obtains:
\[ P_k = 0.25 \left[ I + (-1)^{(k-1)} \Gamma + j^{(k-1)} \left[ F + (-1)^{(k-1)} F^* \right] \right]. \] (18)

Therefore the final expressions required for computing the orthogonal projection matrices are given by:
\[ P_1 = 0.25 \left[ I + \Gamma + 2 \text{Real}(F) \right], \] (19)
\[ P_2 = 0.25 \left[ I - \Gamma - 2 \text{Imaginary}(F) \right], \] (20)
\[ P_3 = 0.25 \left[ I + \Gamma - 2 \text{Real}(F) \right], \] (21)
\[ P_4 = 0.25 \left[ I - \Gamma + 2 \text{Imaginary}(F) \right]. \] (22)

**Method B**

According to a corollary of the spectral theorem, each orthogonal projection matrix \( P_k \) can be expressed as [16, p. 434]:
\[ P_k = g_k(F), \quad k = 1, \cdots, 4 \] (23)

where \( g_k(F) \) is a polynomial in \( F \) which satisfies the conditions:
\[ g_k(\lambda_i) = \delta_{i,k}, \quad i = 1, \cdots, 4. \] (24)

Since matrix \( F \) is diagonalizable and has only 4 distinct eigenvalues, the polynomial \( g_k(F) \) is of the third degree. Using the Lagrange interpolation formula, \( g_k(\lambda) \) can be expressed as:
\[ g_k(\lambda) = \sum_{i=1}^{4} g_k(\lambda_i)f_i(\lambda) \] (25)

---

6 Although matrix \( F \) is symmetric it is not Hermitian.
where
\[ f_i(\lambda) = \prod_{r \neq i} \frac{\lambda - \lambda_r}{\lambda_i - \lambda_r}, \quad i = 1, \ldots, 4. \] (26)

Combining (24)-(26), one gets:
\[ g_k(\lambda) = f_k(\lambda) = \prod_{r \neq k} \frac{\lambda - \lambda_r}{\lambda_k - \lambda_r}. \] (27)

Upon substituting (3) in the above equation, one obtains:
\[ g_1(\lambda) = 0.25\left(1 + \lambda + \lambda^2 + \lambda^3\right), \] (28)
\[ g_2(\lambda) = 0.25\left(1 + j\lambda - \lambda^2 - j\lambda^3\right), \] (29)
\[ g_3(\lambda) = 0.25\left(1 - \lambda + \lambda^2 - \lambda^3\right), \] (30)
\[ g_4(\lambda) = 0.25\left(1 - j\lambda - \lambda^2 + j\lambda^3\right). \] (31)

Using (23) and (28)-(31), one gets the following expressions for the projection matrices:
\[ P_1 = 0.25\left(I + F + F^2 + F^3\right), \] (32)
\[ P_2 = 0.25\left(I + jF - F^2 - jF^3\right), \] (33)
\[ P_3 = 0.25\left(I - F + F^2 - F^3\right), \] (34)
\[ P_4 = 0.25\left(I - jF - F^2 + jF^3\right). \] (35)

Using (3), it immediately follows that the above four expressions are identical to expression (12) obtained using the first approach.

### III. INITIAL ORTHONORMAL EIGENVECTORS OF MATRIX F

The N th order square orthogonal projection matrix \( P_k \) has rank \( r_k \) which is the dimension of the k th eigenspace of matrix F given by Table 1 [12]. The objective here is to compute orthonormal basis for each eigenspace. The singular value decomposition technique will be shown to be the right technique to apply. The singular value decomposition of an arbitrary square matrix A of order N is [18]:
\[ A = U\Sigma V^+ \] (36)
where \( U \) and \( V \) are unitary matrices of order N and 
\[ \Sigma = \text{Diag}\{\sigma_1, \ldots, \sigma_N\}. \] (37)
In the above equation the singular values $\sigma_i, i = 1, \cdots N$ are real and satisfy $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N \geq 0$.

**Lemma 1:**

Let $A$ be a square Hermitian matrix of order $N$ having a modal matrix $V$ and eigenvalues $\lambda_1, \cdots, \lambda_N$ arranged such that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|$. The singular value decomposition of $A$ is given by (36) where

$U = VS$, \hspace{1cm} (38)

$\Sigma = \text{Diag} \{ |\lambda_1|, \cdots, |\lambda_N| \}$, \hspace{1cm} (39)

$S = \text{Diag} \{ s_1, \cdots, s_N \}$ \hspace{1cm} (40)

and

$s_n = \begin{cases} 1 & \text{if } \lambda_n \geq 0 \\ -1 & \text{if } \lambda_n < 0 \end{cases}, \quad n = 1, \cdots, N$. \hspace{1cm} (41)

**Proof:** See Appendix A.

The above lemma implies that for a Hermitian matrix the singular values are equal to the absolute values of its eigenvalues and the right singular vectors are equal to its orthonormal eigenvectors.

If the rank of $A$ is $r$, the matrices $V$ and $\Sigma$ in (36) can be partitioned as:

$V = (V_a \hspace{0.5cm} V_b)$, \hspace{1cm} (42)

$\Sigma = \begin{pmatrix} \Sigma_a & O \\ O & O \end{pmatrix}$ \hspace{1cm} (43)

where

$\Sigma_a = \text{Diag} \{ \sigma_1, \cdots, \sigma_r \}$ \hspace{1cm} (44)

and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. In (42) the $N \times r$ submatrix $V_a$ has orthonormal columns by the unitarity of $V$.

For a Hermitian matrix $A$ of rank $r$, combining (36),(38),(40),(42) and (43), one obtains:

$A = V_a S_a \Sigma_a V_a^+$ \hspace{1cm} (45)

where $S_a$ is the leading diagonal block of order $r$ of matrix $S$ defined by (40).

In the particular case of an orthogonal projection matrix $A$, all eigenvalues are either 1 or 0 [18]. Consequently the diagonal matrices $\Sigma_a$ and $S_a$ in (45) reduce to the identity matrix and (45) simplifies to:

$A = V_a V_a^+$. \hspace{1cm} (46)
Therefore in order to find orthonormal basis for any space of dimension $r$ given its orthogonal projection matrix $A$, one has only to apply the singular value decomposition technique (36) and select the first $r$ columns of matrix $V$ in (42). It should be emphasized that one should never try to get the same result computationally by applying an eigenvalue decomposition procedure since $A$ might have repeated eigenvalues and the corresponding eigenvectors – as evaluated by that procedure – will not generally be orthogonal.

Applying the result (46) to the four projection matrices of last section, one gets:

$$P_k = V_k V_k^+, \quad k = 1, \cdots, 4.$$ (47)

The orthonormal basis of the $k$th eigenspace of matrix $F$ given by the columns of $V_k$ will be taken as initial orthonormal eigenvectors of $F$ corresponding to $\lambda_k$. They will be utilized for deriving the desired Hermite-Gaussian like orthonormal eigenvectors directly needed for defining a discrete fractional Fourier transform that approximates its continuous counterpart.

**IV. HERMITE-GAUSSIAN LIKE EIGENVECTORS**

By sampling the Hermite-Gaussian functions, Pei et. al. obtained approximate eigenvectors for the DFT matrix $F$ that correspond to the exact eigenvalues $\lambda_k$ of (3) as delineated in [4,5]. Those vectors are grouped to form the columns of the four $N \times r_k$ matrices $U_k, k = 1, \cdots, 4$ where the dimensions of the corresponding eigenspaces $r_k$ are given in Table 1. Being approximate rather than exact eigenvectors, the columns of $U_k$ do not belong to the eigenspace $E_k$ corresponding to the eigenvalue $\lambda_k$. The objective here is to find orthonormal basis for $E_k$ to form the columns of a matrix $\hat{U}_k$ that are as close as possible to the columns of matrix $U_k$. Towards achieving that goal, two techniques have been proposed by Pei et. al. [5], namely the Gram-Schmidt Algorithm (GSA) and the Orthogonal Procrustes Algorithm (OPA). A third technique to be termed the Sequential Orthogonal Procrustes Algorithm (SOPA) will be proposed in this paper. In preparation for presenting the new technique, the first two techniques will be surveyed below and cast in a form that will facilitate the comparison. More importantly, they will be proved to produce an output that is invariant under the change of the initial orthonormal basis of $E_k$.

Since the eigenspaces $E_k, k = 1, \cdots, 4$ are orthogonal to each other due to the unitarity of matrix $F$, each eigenspace will be dealt with separately. In order to simplify the notation the
subscript $k$ will be dropped in the remainder of this paper. The $N \times r_k$ matrices $V_k, U_k, \hat{U}_k$ will be written as the $N \times r$ matrices $V, U, \hat{U}$ respectively. The space $E_k$ will be denoted by $E$.

(A) The Gram-Schmidt Algorithm (GSA)

First matrix $U$ will be expressed as the partitioned matrix:

$$U = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix}.$$  \hfill (48)

Each column $u_n$ of $U$ will be projected on $E$ to get $\tilde{u}_n$. Since the resulting $r$ vectors are not orthogonal, they will be orthonormalized by applying the Gram-Schmidt technique in order to get $\hat{u}_n, n = 1, \cdots, r$. More specifically, since the space $E$ is spanned by the orthonormal columns of $V$, vector $\tilde{u}_n$ can be expressed as:

$$\tilde{u}_n = \sum_{m=1}^{\infty} (v_m, u_n) v_m, \quad n = 1, \cdots, r.$$  \hfill (49)

The above equation can be rewritten as:

$$\tilde{u}_n = (v_1, \cdots, v_r) \left[ \begin{array}{c} \langle v_1, u_n \rangle \\ \vdots \\ \langle v_r, u_n \rangle \end{array} \right].$$  \hfill (50)

By virtue of the definition of matrix $V$, the same equation can be compactly expressed as:

$$\tilde{u}_n = V V^* u_n, \quad n = 1, \cdots, r.$$  \hfill (51)

Upon defining the $N \times r$ matrix $\tilde{U}$ as:

$$\tilde{U} = \begin{pmatrix} \tilde{u}_1 & \cdots & \tilde{u}_r \end{pmatrix},$$  \hfill (52)

the $r$ vector equations (51) can be combinedly expressed as:

$$\tilde{U} = V V^* U.$$  \hfill (53)

The Gram-Schmidt technique is next applied to orthonormalize the columns of matrix $\tilde{U}$ in order to get the columns of matrix $\hat{U}$ using the following steps:

1) $\hat{u}_1 = \frac{\tilde{u}_1}{\|\tilde{u}_1\|}$.  \hfill (54)

2) For $s = 2, \cdots, r$:

   a) $y_s = \tilde{u}_s - \sum_{m=1}^{s-1} \langle \hat{u}_m, \tilde{u}_s \rangle \hat{u}_m$.  \hfill (55)
b) \( \hat{u}_s = \frac{y_s}{\|y_s\|} \). 

Matrix \( \hat{U} \) is next defined as:
\[
\hat{U} = (\hat{u}_1 \cdots \hat{u}_r).
\] (57)

Lemma 2:
The result of the Gram-Schmidt Algorithm is invariant under the change of the initial orthonormal basis of the space \( E \) given by the columns of matrix \( V \).

Proof:
Consider a second set of initial orthonormal basis given by the columns of a matrix \( W \) defined by:
\[
W = (w_1 \cdots w_r).
\] (58)

Since the columns of \( V \) form a basis of \( E \), the columns of \( W \) can be expressed as linear combinations of those of \( V \), i.e.,
\[
w_n = \sum_{m=1}^r v_m \alpha_{mn}, \quad n = 1, \ldots, r.
\] (59)

The above vector equations can be compactly expressed as a matrix equation:
\[
W = VG
\] (60)
where \( G \) is a square matrix of order \( r \). It follows immediately that:
\[
W^+W = G^+(V^+V)G.
\] (61)

By the orthonormality of the columns of \( V \) and \( W \) individually, the above equation implies that:
\[
G^+G = I.
\] (62)

Based on this unitarity property of \( G \), it follows from (60) that:
\[
WW^+ = VGG^+V^+ = VV^+.
\] (63)

This proves that the matrix product \( VV^+ \) is invariant under the change of the initial orthonormal basis of the space \( E \). It follows from (53) that \( \hat{U} \) is invariant under the change of \( V \). The same applies to \( \hat{U} \) as can be concluded from (54)-(57).

(Q.E.D.)

(B) The Orthogonal Procrustes Algorithm (OPA)
Here the desired matrix $\hat{U}$ of the Hermite-Gaussian like eigenvectors defined by (57) will be expressed as\(^7\):

$$\hat{U} = VQ$$  \hspace{1cm} (64)

where $Q$ is a unitary matrix to be evaluated such that the square of the Frobenius norm $\|U - \hat{U}\|_F$ is minimized. The solution of this problem is given by the Orthogonal Procrustes Algorithm expounded in [14] and used in [5] and summarized in the following three steps:

1. Form matrix $C$:
   $$C = V^+ U.$$  \hspace{1cm} (65)

2. Find the singular value decomposition of $C$:
   $$C = ADB^+.$$  \hspace{1cm} (66)

3. Compute matrix $Q$:
   $$Q = AB^+.$$  \hspace{1cm} (67)

**Lemma 3:**

The matrix $\hat{U}$ determined by the OPA is invariant under the change of the initial orthonormal basis of the space $E$ given by the columns of matrix $V$.

**Proof:**

Let $V_1$ and $V_2$ be two initial orthonormal bases of $E$. They should be related by:

$$V_2 = V_1 G$$  \hspace{1cm} (68)

where $G$ is a unitary matrix. (This follows along the same lines of (59)-(62)). Let $Q_1$ and $Q_2$ be the corresponding unitary matrices evaluated by the OPA and $\hat{U}_1$ and $\hat{U}_2$ be the resulting optimal matrices given by (64). It follows from (65)-(67) and (64) that:

$$C_i = V_i^* U, \quad i = 1,2$$  \hspace{1cm} (69)

$$C_i = A_i D_i B_i^+, \quad i = 1,2$$  \hspace{1cm} (70)

$$Q_i = A_i B_i^+, \quad i = 1,2$$  \hspace{1cm} (71)

$$\hat{U}_i = V_i Q_i, \quad i = 1,2$$  \hspace{1cm} (72)

Our task reduces to showing that $\hat{U}_2 = \hat{U}_1$. From (68) and (69), one gets:

---

\(^7\) It should be mentioned that in [5], the OPA was erroneously applied since $\hat{U}$ was taken as $\hat{U} = QV$ rather than as $\hat{U} = VQ$. 

---

12
\[ C_2 = V_2^*U = G^*\left(V_1^*U\right) = G^*C_1. \]  \hfill (73)

Substituting (70) in the above equation, one obtains:

\[ C_2 = A_2D_2B_2^* = G^*A_1D_1B_1^*. \]  \hfill (74)

By the uniqueness of the singular values of \( C_2 \), the above equation implies that:

\[ D_2 = D_1. \]  \hfill (75)

By the unitarity of the matrices \( A_1, B_1, A_2, B_2, G \), equation (74) leads to:

\[ D_2 = A_2^*G^*A_1D_1B_1^*B_2. \]  \hfill (76)

Since the diagonal matrices \( D_2 \) and \( D_1 \) have real nonnegative diagonal elements arranged in decreasing order of absolute value, a direct application of Lemma 1 results in:

\[ D_2 = TD_1T^+. \]  \hfill (77)

By comparing (76) and (77), one gets:

\[ A_2^*G^*A_1 = \left(B_1^*B_2\right)^+. \]  \hfill (78)

It follows immediately that:

\[ A_2B_2^* = G^*A_1B_1^*. \]  \hfill (79)

Upon utilizing (71), the above equation reduces to:

\[ Q_2 = G^*Q_1. \]  \hfill (80)

Substituting (68) and (80) in (72), one eventually obtains:

\[ \hat{U}_2 = V_2Q_2 = (V_1G)(G^*Q_1) = V_1Q_1 = \hat{U}_1. \]  \hfill (81)

\( \text{(Q.E.D.)} \)

\( \text{(C) The Sequential Orthogonal Procrustes Algorithm (SOPA)} \)

Although the desired matrix \( \hat{U} \) will be expressed as in (64), the unitary matrix \( Q \) will not be evaluated by minimizing the square of the Frobenius norm \( \| U - \hat{U} \|_F \) as in the OPA but rather its columns \( q_s, s = 1, \cdots, r \) will be evaluated sequentially in the manner to be next delineated. By virtue of the definition of the Frobenius norm of a matrix and the Euclidean norm of a vector [18] and using (48), (57) and (64), one obtains:

\[ J = \sum_{s=1}^r J_s \]  \hfill (82)

where

\[ J = \| U - \hat{U} \|_F^2 = \| U - VQ \|_F^2, \]  \hfill (83)
\[ J_s = \|u_s - \hat{u}_s\|_2^2 = \|u_s - Vq_s\|_2^2, \quad (84) \]
\[ Q = (q_1 \cdots q_r). \quad (85) \]

In the OPA, matrix \( Q \) has been evaluated by minimizing the total performance index \( J \) of (83). In the SOPA there will be \( r \) stages of sequential minimizations. In stage \( s \), the column \( q_s \) of \( Q \) will be evaluated by minimizing the partial performance index \( J_s \) of (84) subject to the constraints that \( q_s \) will be orthogonal to the previously evaluated columns \( q_k, k = 1, \cdots, s - 1 \) and be of unit norm in order to satisfy the unitarity of \( Q \). These \( s \) constraints can be expressed as:
\[ q_s^* q_s = 0 \quad , \quad k = 1, \cdots, s - 1 \quad (86) \]
\[ q_s^* q_s = 1. \quad (87) \]

The first set of \((s - 1)\) orthogonality constraints can be compactly expressed as:
\[ C_{s-1} q_s = 0 \quad (88) \]
where
\[ C_{s-1} = \begin{pmatrix} q_1^* \\ \vdots \\ q_{s-1}^* \end{pmatrix} \quad (89) \]

It should be mentioned that the rows of the \((s - 1) \times r\) matrix \( C_{s-1} \) are linearly independent because of being orthogonal due to the way they have been sequentially generated. For mathematical tractability, one will set aside the quadratic constraint (87) for a while and minimize \( J_s \) subject to the linear constraints (88) and call the resulting vector \( x_s \). Next by normalizing \( x_s \) in order to satisfy the normalization condition (87), one obtains \( q_s \). More specifically, the constrained minimization problem defined by the quadratic criterion (84) and the linear constraints (88) is solved in Appendix B and its solution is given by:
\[ x_s = (V^* V)^{-1} \left( I - C_{s-1}^* \left[ C_{s-1} (V^* V)^{-1} C_{s-1}^* \right] I C_{s-1} (V^* V)^{-1} \right) V^* u_s. \quad (90) \]

By the orthonormality of the columns of matrix \( V \), the above equation reduces to:
\[ x_s = \left( I - C_{s-1}^* C_{s-1} \right) V^* u_s. \quad (91) \]

By utilizing the orthonormality of the rows of matrix \( C_{s-1} \) of (89), the above result simplifies to:
\[ x_s = \left( I - C_{s-1}^* C_{s-1} \right) V^* u_s \quad , \quad s = 2, \cdots, r. \quad (92) \]
In the special case of generating $x_1$, one has the unconstrained minimization problem defined by criterion $J_1$ of (84) whose solution is given by:

$$x_1 = V^+ u_1.$$  \hfill (93)

The quadratic constraint (87) accounting for the normalization condition will be satisfied by computing:

$$q_s = \frac{1}{\|x_s\|} x_s, \quad s = 1, \ldots, r.$$  \hfill (94)

Therefore the sequential orthogonal procrustes algorithm can be summarized in the following steps:

1) For $s = 1$:
   a) $x_1 = V^+ u_1$
   b) $q_1 = \frac{1}{\|x_1\|} x_1$
   c) $Q = q_1$
   d) $C = [\ ]$ (the null matrix)

2) For $s = 2, \ldots, r$:
   a) Augment matrix $C$ by the row vector $q^+_{s-1}$
   b) $x_s = (I - C^* C)^{1/2} \tilde{V}^* u_s$
   c) $q_s = \frac{1}{\|x_s\|} x_s$
   d) Augment matrix $Q$ by the column vector $q_s$

3) Generate $\tilde{U}$ according to (64).

**Lemma 4:**

The matrix $\tilde{U}$ determined by the SOPA is invariant under the change of the initial orthonormal basis of the space $E$ given by the columns of matrix $V$.

**Proof:**

Let $V_1$ and $V_2$ be two initial orthonormal bases of $E$ which should be related by (68). Let $Q_i$ be the $Q$ matrix determined by the SOPA corresponding to $V_i$ and let $\tilde{U}_i$ be the resulting matrix $\tilde{U}$. An extra subscript $i$ will be introduced in the vectors $x_i$ and $q_i$ and matrix $C_{s-1}$.  

15
to become \( x_{i,t}, q_{i,t} \) and \( C_{i,t-1} \) respectively when they are computed based on matrix \( V_i \). From (93) and (68), one obtains:

\[
x_{2,1} = V_2^+ u_1 = G^+ V_1^+ u_1 = G^+ x_{1,1} .
\]

(95)

By virtue of (94) and the unitarity of \( G \), one gets:

\[
q_{2,1} = G^+ q_{1,1} .
\]

(96)

From (89) and (96), it follows that:

\[
C_{2,1} = q_{2,1}^* G = C_{1,1} G .
\]

(97)

From (92), (97) and (68), one obtains:

\[
x_{2,2} = \left( I - C_{2,1}^* C_{2,1} \right) V_2^+ u_2 = \left( I - G^* C_{1,1}^* C_{1,1} G \right) G^+ V_1^+ u_2 = G^* \left( I - C_{1,1}^* C_{1,1} \right) V_1^+ u_2.
\]

(98)

It follows from the above equation and (94) that:

\[
q_{2,2} = G^+ q_{1,2} .
\]

(99)

By the same token, it can be shown that:

\[
q_{2,s} = G^+ q_{1,s} , \quad s = 1, \ldots, r .
\]

(100)

From (85) and (100), one obtains:

\[
Q_2 = G^+ Q_1 .
\]

(101)

The above equation together with (64) and (68) lead to:

\[
\hat{U}_2 = V_2 Q_2 = \left( V_1 G \right) (G^+ Q_1) = V_1 Q_1 = \hat{U}_1 .
\]

(102)

(Q.E.D.)

V. THE EQUALITY OF THE OUTPUTS OF THE GSA AND SOPA

In the GSA, one starts by projecting the vectors \( u_s, s = 1, \ldots, r \) on the column space of matrix \( V \) to get the nonorthogonal vectors \( \tilde{u}_s, s = 1, \ldots, r \). Next by applying the Gram-Schmidt orthonormalization technique, one sequentially obtains the orthonormal vectors \( \hat{u}_s, s = 1, \ldots, r \). In the SOPA, one sequentially obtains the columns \( q_s, s = 1, \ldots, r \) of the unitary matrix \( Q \). This corresponds to the sequential evaluation of the vectors \( \hat{u}_s, s = 1, \ldots, r \) since (64), (57) and (85) imply that:

\[
\hat{u}_s = V q_s , \quad s = 1, \ldots, r .
\]

(103)
It will be shown below that the GSA and SOPA produce identical results. Towards that goal one starts by a further manipulation of the equations pertaining to both techniques only for the sake of proving the equality of their outputs.

(A) The GSA

Manipulating (55), one obtains:

\[ y_s = \tilde{u}_s - \sum_{m=1}^{s-1} \tilde{u}_m \left( \tilde{u}_m^+ \tilde{u}_s \right) \]

\[ = \tilde{u}_s - \left( \sum_{m=1}^{s-1} \tilde{u}_m \tilde{u}_m^+ \right) \tilde{u}_s \]

\[ = \tilde{u}_s - \hat{U}_{s-1}^+ \hat{U}_{s-1} \tilde{u}_s \]

where

\[ \hat{U}_{s-1} = (\tilde{u}_1 \ \cdots \ \tilde{u}_{s-1}) \]

Substituting (51) in (104), one gets:

\[ y_s = VV^+ u_s - \hat{U}_{s-1}^+ \hat{U}_{s-1} VV^+ u_s \]

\[ = (A - B_{s-1} A) u_s \]

where

\[ A = VV^+ , \]

\[ B_{s-1} = \hat{U}_{s-1}^+ \hat{U}_{s-1} \] (108)

Since the vectors \( \tilde{u}_s \), \( s = 1, \cdots, r \) lie in the column space of matrix \( V \), it follows from (105) that

\[ \hat{U}_{s-1} = VW \]

where \( W \) is an \( r \times (s-1) \) matrix. Consequently

\[ \hat{U}_{s-1}^+ \hat{U}_{s-1} = W^+ (V^+V) W . \] (110)

By the orthonormality of the columns of \( \hat{U}_{s-1} \) and those of \( V \), the above equation results in:

\[ W^+ W = I . \] (111)

From (107), (108) and (109), one obtains:

\[ B_{s-1} A = VWW^+ (V^+ V) V^+ = (VW) (W^+ V^+) = \hat{U}_{s-1}^+ \hat{U}_{s-1} = B_{s-1} . \] (112)

Therefore (106) simplifies to:

\[ y_s = (A - B_{s-1}) u_s \quad s = 2, \cdots, r . \] (113)
In the special case of \( s = 1 \), (51) and (107) result in:
\[
y_1 = \tilde{u}_1 = VV^* u_1 = Au_1 .
\] (114)

(B) The SOPA

One starts by defining:
\[
z_s = Vx_s , \quad s = 1, \ldots , r
\] (115)
where the vectors \( x_s , s = 1, \ldots , r \) are given by (92) and (93). The orthonormality of the columns of \( V \) implies that:
\[
\|z_s\| = \|x_s\| , \quad s = 1, \ldots , r .
\] (116)

Equations (115), (116), (94) and (103) result in:
\[
\frac{z_s}{\|z_s\|} = Vq_s = \hat{u}_s , \quad s = 1, \ldots , r .
\] (117)

Since (94) implies that \( x_s \) is an unnormalized version of \( q_s \), the above equation implies that \( z_s \) defined by (115) can be interpreted as an unnormalized version of \( \hat{u}_s \). Combining (115) and (92), one obtains:
\[
z_s = \left[ VV^* - VC_{s-1}^+ \right] u_s , \quad s = 2, \ldots , r .
\] (118)

Using (89), (103) and (105), it follows that:
\[
VC_{s-1}^+ = \left[ \begin{array}{ccc} q_1 & \cdots & q_{s-1} \end{array} \right] \left( \begin{array}{ccc} \hat{u}_1 & \cdots & \hat{u}_{s-1} \end{array} \right) = \hat{U}_{s-1} .
\] (119)

Substituting the above equation in (118) and utilizing (107) and (108), one gets:
\[
z_s = \left[ VV^* - \hat{U}_{s-1} \hat{U}_{s-1}^+ \right] u_s = (A - B_{s-1}) u_s , \quad s = 2, \ldots , r .
\] (120)

In the special case of \( s = 1 \), it follows from (115), (93) and (107) that:
\[
z_1 = Vx_1 = VV^* u_1 = Au_1 .
\] (121)

(C) The identity

From (107) and (108), it is obvious that matrix \( A \) is the same for both the GSA and the SOPA since it is an input matrix while the matrices \( B_{s-1} , s = 2, \cdots , r \) are computed for each algorithm separately. From (114) and (121) it is clear that \( y_1 = z_1 \); and from (56) and (117) it follows that \( \hat{u}_1 \) is the same for both algorithms. By virtue of (105) and (108), one concludes that \( \hat{U}_1 \) and \( B_1 \) are identical for both algorithms. Equations (113) and (120) imply that \( y_2 = z_2 \) and consequently \( \hat{u}_2 \) are the same for both algorithms.
manner, one concludes that both the GSA and SOPA result in exactly the same set of vectors \( \hat{u}_s, s = 1, \cdots, r \).

VI. SIMULATION RESULTS

Orthonormal eigenvectors of the DFT matrix \( F \) have been computed using the following three techniques:

1) The P method where one only obtains initial orthonormal vectors by the singular value decomposition of the projection matrices \( P_k, k = 1, \cdots, A \) of \( F \) according to (47) as was explained in section III.

2) The Orthogonal Procrustes Algorithm (OPA) explained in section IV (B).

3) The Sequential Orthogonal Procrustes Algorithm (SOPA) explained in section IV (C).

It should be mentioned that even if one intends to apply a refined technique such as the second or third one, he should start by applying the first technique in order to generate initial orthonormal eigenvectors of \( F \) to be used as input to the advanced technique. Since the main goal is to generate Hermite-Gaussian like orthonormal eigenvectors of \( F \), the error vectors between the eigenvectors of \( F \) and samples of the Hermite-Gaussian functions of the same order have been computed as it was done in [5]. The norms of these error vectors have been plotted versus \( k \) for \( N_k = 1, \cdots, L \) where \( k \) denotes the columns of the modal matrix of \( F \). Figures 1 and 2 show the results for \( N = 64 \) and 128 respectively. Since in the P method the act of approximating the Hermite-Gaussian functions was not taken into account, it is quite expected that the P method has the largest error among the three techniques being compared. For the OPA or SOPA the error tends to increase on the average with \( k \). The interpretation is that the samples of the Hermite-Gaussian functions are approximate eigenvectors of \( F \) with an approximation error that grows with the order of those functions [4,5]. Consequently the error between the exact Hermite-Gaussian like eigenvectors determined by the OPA or SOPA and the approximate eigenvectors tends to increase with \( k \). Upon comparing the OPA and SOPA, one clearly notices that the OPA has a lower rate of growth of the error. The interpretation is that for the SOPA the number of linear constraints expressed by (88) and (89) increases with the order thus restricting the freedom left in the solution space for minimizing the criterion \( J_s \) given by (84). This is to be contrasted to the OPA where matrix \( \hat{U} \) rather than its individual columns is batch evaluated. On the other hand, the SOPA has the merit that the error begins to be noticeable at a value of \( k \) larger than that for the OPA.
VII. CONCLUSION

A new technique has been developed for generating *initial* orthonormal eigenvectors of the Discrete Fourier Transform matrix $F$ based on the singular value decomposition of the projection matrices of $F$ on its eigenspaces after deriving efficiently computable expressions for these projection matrices. In order to generate Hermite-Gaussian like eigenvectors of $F$ given the *initial* ones, a new method called the Sequential Orthogonal Procrustes Algorithm (SOPA) has been proposed based on the sequential evaluation of the columns of a unitary matrix rather than the batch evaluation of that matrix as in the Orthogonal Procrustes Algorithm (OPA). Surprisingly the output of the SOPA has been proved to be equal to that of the Gram-Schmidt Algorithm (GSA). Furthermore It has been proved that for any of the GSA, OPA or SOPA, the output is invariant under the change of the input *initial* orthonormal eigenvectors of $F$.
APPENDIX A

Proof of Lemma 1

The modal decomposition of a Hermitian matrix $A$ is:

\[ A = V \Lambda V^+ \]  \hspace{1cm} (A1)

where

\[ \Lambda = \text{Diag}\{\lambda_1, \ldots, \lambda_N\} \]  \hspace{1cm} (A2)

and all eigenvalues $\lambda_n$, $n = 1, \cdots, N$ are real [15]. The diagonal matrix $\Lambda$ can be expressed as:

\[ \Lambda = S \Sigma \]  \hspace{1cm} (A3)

where the real diagonal matrices $\Sigma$ and $S$ are defined by (39)-(41). Substituting (A3) in (A1), one gets:

\[ A = (V S) \Sigma V^+ . \]  \hspace{1cm} (A4)

By comparing (36) and (A4), one obtains (38).

(Q.E.D.)
APPENDIX B

Statement of the problem:
Find the r-dimensional vector $x$ that minimizes:

$$J = \| u - Vx \|^2$$

subject to the constraints:

$$Cx = 0$$

where $u$ is an N-dimensional vector, $V$ is an $N \times r$ matrix with linearly independent columns, $C$ is an $(s-1) \times r$ matrix with linearly independent rows and $(s-1) < r < N$.

Solution:
Augmenting the constraints (B2) to the criterion (B1) by means of the complex vector $\lambda$ of Lagrange multipliers, one gets the following real augmented criterion:

$$J_a = J - \lambda^* Cx - (Cx)^* \lambda .$$

By virtue of the definition of the Euclidean norm, one obtains from (B1) and (B3):

$$J_a = u^* u + x^* V^* Vx - u^* Vx - x^* V^* u - \lambda^* Cx - x^* C^* \lambda .$$

Since $J_a$ is a real-valued scalar function of the complex vector $x$ and its complex conjugate $x^*$, a necessary and sufficient condition for minimization is:

$$\nabla_{x^*} J_a = 0$$

where in finding the gradient vector, one should view $x$ and $x^*$ as two different vectors, i.e. one should treat $x$ as a constant vector when evaluating $\nabla_{x^*}$. [19]. Consequently it follows from (B4) that:

$$\nabla_{x^*} J_a = V^* Vx - V^* u - C^* \lambda .$$

Upon applying condition (B5), one gets:

$$x = (V^* V)^{-1} (V^* u + C^* \lambda) .$$

By applying condition (B2) in order to evaluate vector $\lambda$, one obtains:

$$\lambda = -\left[ C (V^* V)^{-1} C^* \right]^{-1} C (V^* V)^{-1} V^* u .$$

Substituting (B8) in (B7), one obtains:

$$x = (V^* V)^{-1} \left\{ I - C^* \left[ C (V^* V)^{-1} C^* \right]^{-1} C (V^* V)^{-1} \right\} V^* u .$$
REFERENCES


Table 1: The dimensions $r_k$ of the four eigenspaces

<table>
<thead>
<tr>
<th>$\lambda_k$</th>
<th>$\lambda_1$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>4m</td>
<td>m+1</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m-1</td>
</tr>
<tr>
<td>4m+1</td>
<td>m+1</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>4m+2</td>
<td>m+1</td>
<td>m</td>
<td>m+1</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>4m+3</td>
<td>m+1</td>
<td>m+1</td>
<td>m+1</td>
<td>m</td>
<td>m</td>
</tr>
</tbody>
</table>
Fig. 1: Norm of the error vectors between the exact and approximate eigenvectors for $N = 64$.

Fig. 2: Norm of the error vectors between the exact and approximate eigenvectors for $N = 128$. 