

A COMPARATIVE STUDY OF COMMUTING MATRIX APPROACHES FOR THE DISCRETE FRACTIONAL FOURIER TRANSFORM

Ishwor Bhatta and Balu Santhanam

Dept. of ECE, University of New Mexico,
Albuquerque, NM: 87131, USA
Tel: 505 277-1611, Fax: 505 277 1439
Email: bsanthan,ibhatta@unm.edu.

ABSTRACT

As an extension of the conventional Fourier transform and as a time-frequency signal analysis tool, the fractional Fourier transforms (FRFT) are suitable for dealing with various types of non-stationary signals. Computation of the discrete fractional Fourier transform (DFRFT) and its chirp concentration properties are both dependent on the basis of DFT eigenvectors used in the computation. Several DFT-eigenvector bases have been proposed for the computation of transform, and there is no common framework for comparing them. In this paper, we compare several different approaches from a conceptual viewpoint and review the differences between them.

We discuss five different approaches to find centered-DFT (CDFT) commuting matrices and the various properties of these commuting matrices. We study the properties of the eigenvalues and eigenvectors of these commuting matrices to determine whether they resemble those of corresponding continuous Gauss-Hermite operator. We also measure the performance of these five approaches in terms of: mainlobe-to-sidelobe ratio, 10-dB bandwidth, quality factor, linearity of eigenvalues, chirp parameter estimation error, and, finally the peak-to-parameter mapping regions. We compare the five approaches using these performance metrics and point out that the modified QMFD approach produces the best results in terms of bandwidth of the spectral peak for a chirp, invertibility of the peak-parameter mapping, linearity of the eigenvalue spectrum and chirp parameter estimation errors.

Index Terms— Discrete fractional Fourier transform, linear chirp signal, chirp parameter estimation, parameter estimation error, peak-to-parameter mapping, invertibility region.

1. INTRODUCTION

Pei, Hsue and Ding used the error-norm parameter to compare the eigenvectors of different DFT commuting matrices [1] to check the similarity between the eigenvector and the continuous Gauss-Hermite(G-H) function. Serbes and Durak-Ata also used the same parameter for comparison in [2]. These analyses only determine how close the eigenvectors are to sampled G-H functions. As described in [3], the linearity of the eigenvalue spectrum is another important parameter used to determine the closeness of the generating matrix to the G-H functions. [4] further introduced the valid mapping region criteria for peak-to-parameter mapping estimation. In addition to the above mentioned parameters, we employed a few other parameters to determine the best among the various CDFT commuting matrix approaches.

Using these seven parameters we seek to determine the best CDFT commuting matrix among the five different commuting matrix methods. The eigenvectors of these commuting matrices are related to the continuous G-H function. Therefore, after this comparison we will be able to find the particular commuting matrix whose eigenvectors and eigenvalues are closest to those of the continuous G-H operator. In addition, the comparison will allow us to measure the quality of the peak that contains information regarding the chirp rate and center frequency parameters.

2. COMPUTING THE DFRFT

The Dickinson and Steiglitz (D-S) Method: Dickinson and Steiglitz defined a DFT commuting matrix [5], whose eigenvectors look like G-H function, as

$$\mathbf{S} = \begin{pmatrix} 2 & 1 & 0 & \dots & 1 \\ 1 & 2 \cos \omega & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 2 \cos(N-1)\omega \end{pmatrix}, \quad (1)$$

where $\omega = \frac{2\pi}{N}$.

This makes the matrix \mathbf{S} commute with the DFT. If we change the range of n from $0 \leq n \leq N-1$ to $|n| \leq \frac{N-1}{2}$, we will have a new \mathbf{S} -matrix which will commute with CDFT.

$$i.e. [\mathbf{W}, \mathbf{S}_{cen}] = \mathbf{W} * \mathbf{S}_{cen} - \mathbf{S}_{cen} * \mathbf{W} = 0, \quad (2)$$

where \mathbf{W} is the Centered DFT and \mathbf{S}_{cen} is the Centered version of \mathbf{S} -matrix Santhanam and Vargas-Rubio focused their attention on the centered version of the DFT matrix operator \mathbf{W} [6].

$$\mathbf{W}_{mn} = \frac{1}{\sqrt{N}} e^{-j \frac{2\pi}{N} (m-a)(n-a)}, \quad (3)$$

where the shift parameter $a = \frac{N-1}{2}$.

The Bilinear Transformation Method: Another approach to generating a DFT commuting matrix is introduced by Serbes and Durak-Ata in [2]. They defined a new DFT commuting matrix as;

$$\mathbf{B} = \mathbf{B}_1^{-1} \mathbf{E}_2 + \mathbf{W} \mathbf{B}_1^{-1} \mathbf{E}_2 \mathbf{W}^{-1}, \quad (4)$$

where

$$\mathbf{B}_1 = \begin{bmatrix} k & 1 & 0 & \dots & \dots & 0 & 1 \\ 1 & k & 1 & \dots & \dots & 0 & 0 \\ 0 & 1 & k & \dots & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & \dots & 1 & k \end{bmatrix}. \quad (5)$$

$$\mathbf{E}_2 = \begin{bmatrix} -2 & 1 & 0 & \dots & \dots & 0 & 1 \\ 1 & -2 & 1 & \dots & \dots & 0 & 0 \\ 0 & 1 & -2 & \dots & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & \dots & 1 & -2 \end{bmatrix}. \quad (6)$$

\mathbf{W} = DFT matrix and $k = 10^9$ is used in this paper.

The Infinite Order Second Derivative Approximation Method: Inspired by the work of Grunbaum [7], Pei, Hsue and Ding proposed another DFT-commuting matrix in [1], whose eigenvectors are even closer to the continuous G-H function than those of the Dickinson-Steiglitz matrix. The matrix proposed by them is given as;

$$\mathbf{M}_{2k} = \sum_{m=1}^k (-1)^{m-1} \frac{2[(m-1)!]^2}{(2m)!} \mathbf{D}^m, \quad (7)$$

where

$$\mathbf{D} = \begin{bmatrix} -2 & 1 & 0 & \dots & \dots & 0 & 1 \\ 1 & -2 & 1 & \dots & \dots & 0 & 0 \\ 0 & 1 & -2 & \dots & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & \dots & 1 & -2 \end{bmatrix} \quad (8)$$

\sim the second order symmetric difference matrix.

For the purpose of this paper, we have used $k = 2$:

$$\mathbf{M}_4 = \frac{-1}{12} \mathbf{D}^2 + \mathbf{D}. \quad (9)$$

The DFT commuting matrix based on this analysis is;

$$\mathbf{S}_4 = \mathbf{M}_4 + \mathbf{W} \mathbf{M}_4 \mathbf{W}^{-1}. \quad (10)$$

We refer to this matrix \mathbf{S}_4 , which commutes with DFT, as a higher-order S-matrix for the purpose of this paper.

The Grunbaum Method: Another approach to obtaining the DFT eigenvectors uses the tri-diagonal commuting matrix introduced by Grunbaum [7]. Mugler and Clary modified the Grunbaum tri-diagonal incorporating a scaling factor, and the resultant eigenvectors very closely resemble the G-H functions [8]. The tri-diagonal commutator of Grunbaum is defined via its diagonal and off-diagonal elements in [8] as;

$$\mathbf{T}_{mn} = \begin{cases} -2 \cos(\pi N \tau) \sin(\pi \mu \tau) \sin(\pi(N - \mu - 1)\tau), & \text{if } m = n, 0 \leq n \leq N - 1 \\ \sin(\pi \mu \tau) \sin(\pi(N - \mu)\tau), & \text{if } m = n + 1, n - 1, 1 \leq n \leq N - 1 \\ 0, & \text{otherwise} \end{cases}, \quad (11)$$

where $0 \leq \mu \leq N - 1$ and $\tau = \frac{1}{N}$.

The QMFD Method: Santhanam et. al. defined a discrete version of the G-H differential operator \mathbf{H} [3] that furnishes the basis for the centered version of the DFT matrix and simultaneously has eigenvalues and eigenvectors that closely resemble those of the continuous G-H operator. The CDFT commuting matrix is defined as;

$$\mathbf{T} = c_1(\mathbf{P}^2 + \mathbf{Q}^2) + c_2 \mathbf{C}_1^H \mathbf{C}_1 + c_3 \mathbf{I}, \quad (12)$$

where

$$c_1 = 1, c_2 = -c_3 = -\frac{\pi^2}{N^2}, \quad N \text{ is the size of DFT matrix,} \quad (13)$$

$$\mathbf{P} = \mathbf{W} \mathbf{Q} \mathbf{W}^H, \quad \mathbf{C}_1 = \mathbf{Q} \mathbf{P} - \mathbf{P} \mathbf{Q} \quad (14)$$

$$\mathbf{Q}_{rr} = q[r] = \sqrt{\frac{2\pi}{N}} r, \quad -\frac{N-1}{2} \leq r \leq \frac{N-1}{2}, \quad (15)$$

and \mathbf{I} is the identity matrix of dimension N .

We changed the range of r that normally spans $r \in [-\frac{N-1}{2}, \frac{N-1}{2}]$ to the zero locations of the N th order G-H function to see improvement in terms of linearity of the eigenvalue spectrum and the invertibility region in the peak-to-parameter mapping, and we found slight improvement in both metrics. Therefore, we retain the values of r as the zero location of the N^{th} order G-H operator throughout the paper, and we referred to this method as 'modified QMFD' or 'QMOD'.

3. RESULTS

While each of the methods produce a peak in the chirp rate versus frequency plane for a chirp signal, the sharpness and the width of the peak varies based on the basis used. In this section, we attempt to quantify this ability by comparing the methods with respect to following metrics. In the ideal case, a chirp would be transformed into a Dirac impulse for a specified center frequency and chirp rate.

Mainlobe-to-Sidelobe Ratio: For the purpose of this paper, the ratio of the peak value to the value of the second peak is defined as the mainlobe-to-sidelobe ratio. We can see those peak values and side-lobes in figure 1. In order to compute mainlobe-to-sidelobe ratio we first took the absolute value of the row of MA-CDFRFT matrix where the peak occurs. We then consider the highest peak as mainlobe and the second peak as sidelobe. Finally, we took the ratio of these two values as mainlobe-to sidelobe ratio (MLSLR). Figure 2 shows the MLSLR for different combinations of c_r and w_c .

From these figures, it is clear that the MLSLR increases as N increases only in the case of the Grunbaum basis. Also, the QMOD method has a better MLSLR in the case of zero central frequency until $N = 512$, whereas the Grunbaum method has a better MLSLR in the case of non-zero central frequency. Therefore, we can conclude that the QMOD method and the Grunbaum method are the best choice among the five methods towards attaining a better MLSLR. We observed that the MLSLR decreases for some methods for higher values of N .

10-dB Bandwidth: Figure 3 shows the 10-dB bandwidth comparison for different combinations of c_r and w_c . From these figures, we can see that the bandwidth continuously decreases only for the QMOD basis in the case of zero central frequency. And for non-zero central frequency, the bandwidth continuously decreases as N increases. The bilinear transformation method and the S-matrix method have the least bandwidth for non-zero central frequency.

Quality Factor: Quality factor, also known as Q-factor, is a dimensionless quantity which characterizes a signal's bandwidth relative to its central frequency. It measures the quality of the peak in relationship with central frequency. Mathematically, it is the ratio of the central frequency to the bandwidth of the signal. i.e.

$$Q - factor = \frac{w_c}{BW}, \quad (16)$$

where w_c = central frequency and BW = bandwidth of the signal.

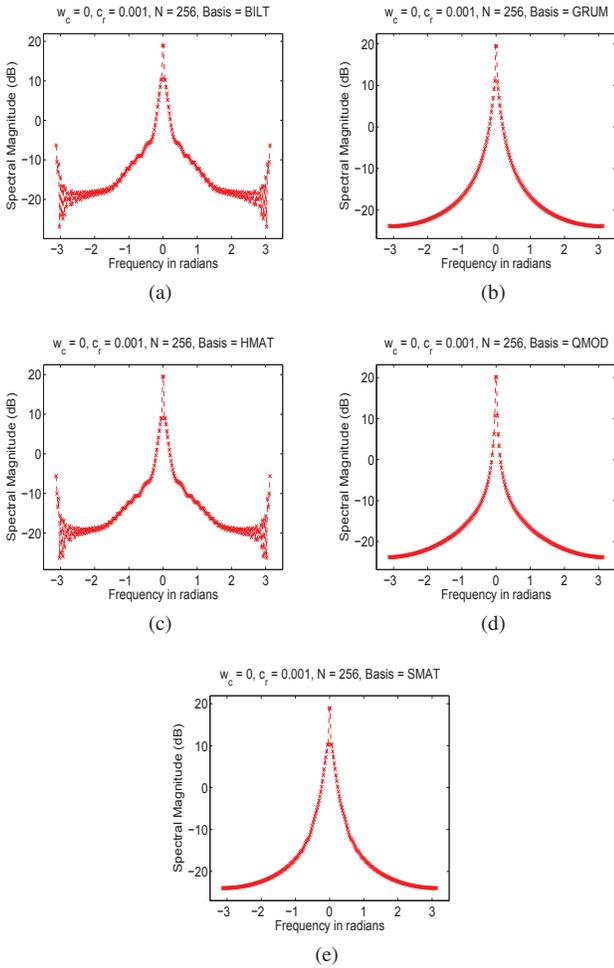


Fig. 1: Plot of slice of MA-CDFRFT at $r = 68$ for $c_r = 0.001$, $w_c = 0$ and $N = 256$ obtained from (a) Bilinear transformation method, (b) Grunbaum method, (c) Higher order S-matrix method, (d) QMOD method and (e) S-matrix method.

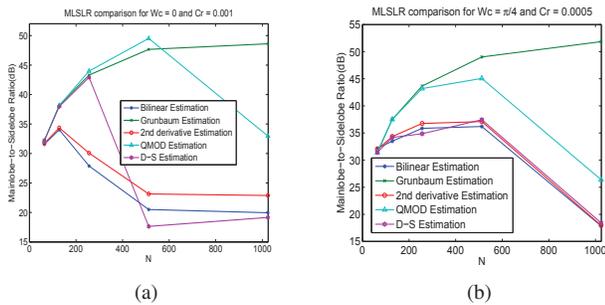


Fig. 2: MLSLR comparison: (a) $c_r = 0.001$ and $w_c = 0$ and (b) $c_r = 0.0005$ and $w_c = \frac{\pi}{4}$.

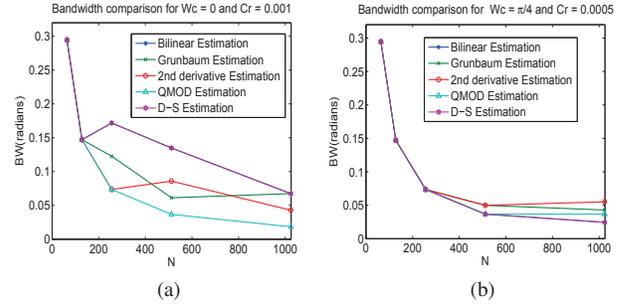


Fig. 3: 10-dB BW comparison for (a) $c_r = 0.001$ and $w_c = 0$ and (b) $c_r = 0.0005$ and $w_c = \frac{\pi}{4}$. The QMOD method exhibits a steady decrease in BW with increase in N for a chirp signal with zero central frequency. For non-zero central frequency, all the five methods exhibit a steady decrease in BW with increase in N , but the S-matrix method and the bilinear transformation method produce best result in terms of BW.

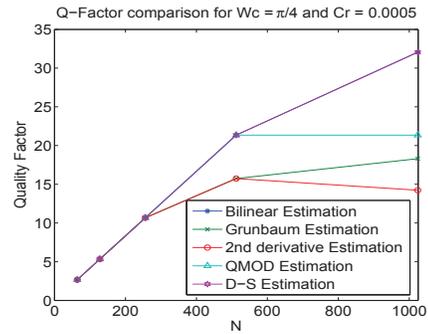


Fig. 4: Q-factor comparison of peak obtained for a chirp with $c_r = 0.0005$ and $w_c = \frac{\pi}{4}$. The S-matrix method and the bilinear transformation method produce the best results in terms of Q-factor.

We set the central frequency as $\frac{\pi}{4}$ and chirp rate as 0.0005 of a chirp signal to find the Q-factor of the peaks for all the five methods. Therefore the signal used to find Q-factor became;

$$x[n] = e^{j(0.0005m^2 + \frac{\pi}{4}n)}, \quad 0 \leq n \leq 255, \quad m = n - \frac{255}{2}. \quad (17)$$

Figure 4 depicts the Q-factor comparison for $c_r = 0.0005$ and $w_c = \frac{\pi}{4}$. From this plot, it is clear that the Q-factor increases for all the bases as N increases, which is obvious from the previous section, because the lower bandwidth requirement for a fixed center frequency is equivalent to a higher quality factor. We can see from the figure that the bilinear transformation method and the S-matrix method are the best choice in terms of Q-factor considerations.

Linearity of the Eigenvalue Spectrum: As described in Santhanam et. al. in [9] the eigenvalues of the G-H operator are linearly spaced. Therefore the eigenvalues of the matrices obtained from all the five different approaches, should be linear in spacing in order to resemble the eigenvalues of the continuous G-H operator. We first plotted the eigenvalues of the five different matrices, to determine the extent of the linearity of eigenvalue spectrum.

Then we calculated the percentage of number of points where the eigenvalues spread linearly for different values of N . Figure 5 shows the percentage of number of points where the eigenvalues are

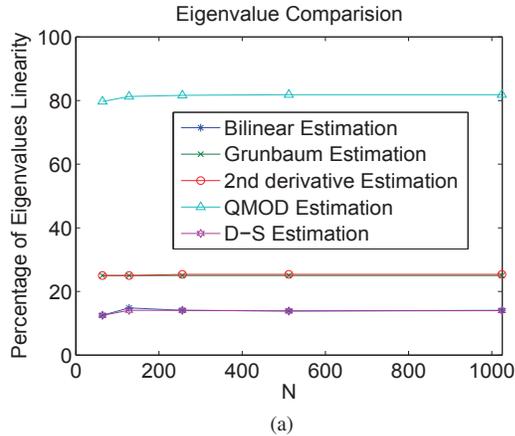


Fig. 5: Percentage of points where eigenvalue spacing is linear for the various commuting matrix approaches.

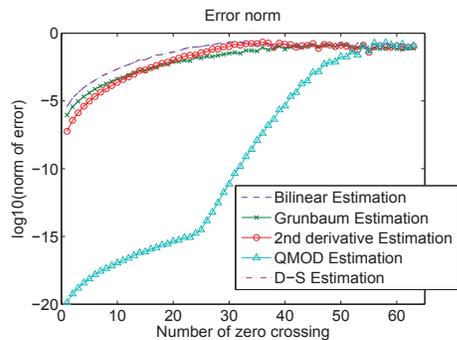


Fig. 6: Error-norm comparison for $N = 64$. The QMOD method exhibits the least error norm for smaller values of zero-crossings.

as linear as that of the continuous G-H operator.

From figure 5, it is seen that the eigenvalue spectrum from the QMOD method best resembles that of the continuous G-H operator. Linearity extends to about 80% of the total points which is far better than all other methods. Furthermore, this linearity steadily increases with increase in N , a fact consistent with the asymptotic convergence of the QMOD matrix to the continuous G-H operator [9].

Error-norm of the Eigenvectors: The eigenvectors of the matrices obtained from the different approaches resemble the corresponding G-H function; however they are not exactly the same. Pei, Hsue and Ding defined the error-norm of eigenvectors as the second norm of the difference between the eigenvectors obtained from the G-H like eigenvector and the samples of its corresponding continuous G-H function [1]. Figure 6 depicts the error-norm comparison for $N = 64$ for all the methods we discussed.

From this figure, it is clear that the error-norm increases as the number of zero crossings increases for all the methods, as we discussed in the previous chapter. We can also see that the QMOD method results in a very low error-norm in comparison to the other methods. Therefore, eigenvectors of the QMOD matrix better resemble the eigenvectors of the corresponding G-H operator in comparison to the other methods from the perspective of error-norm.

Chirp Parameter Estimation Error: Subspace decomposition techniques have been investigated for use in conjunction with the DFRFT with the aim of providing a robust and accurate estimation

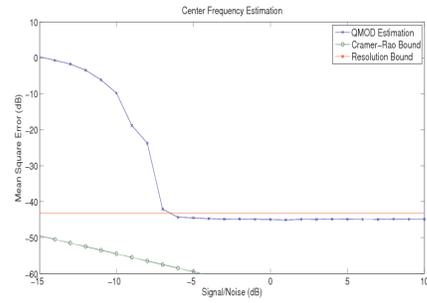
in the presence of noise [10]. Peacock and Santhanam discussed the chirp parameter estimation error using 2D peak picking [4]. We used parameter estimation error as one of the bases for comparison of five proposed approaches. We measured the parameter estimation error of those five approaches and compared them to the Cramer-Rao lower bound and resolution bound. Figure 7 (c), (d) show the parameter estimation error using 2D peak picking for central frequency and chirp rate, for $N = 256$.

From this figure, we can see that the estimation error for the QMOD method attains the resolution bound for higher SNR for both center frequency estimation and chirp rate estimation, and the error for the Grunbaum method is close to that of the QMOD method. However the other three methods result in significantly more parameter estimation error. This is attributable to the fact that the peak-to-parameter mapping depicts multiple disconnected regions of chirp parameters mapping to the same peak location. We can also observe that, both the center frequency estimation and chirp rate estimation errors for the bilinear transformation method do not decrease as gradually as in other methods. This is because of the fact that the particular combination of center frequency and chirp rate, where the mean square error goes up, does not lie within the invertible mapping region. Figure 7 (a), (b) show the parameter estimation error calculated using the cross-hair technique combined with the minimum-norm-subspace technique. From this figure, we can see that the chirp parameter estimation error for the QMOD method approaches the Cramer-Rao lower bound for chirp parameter estimation [11].

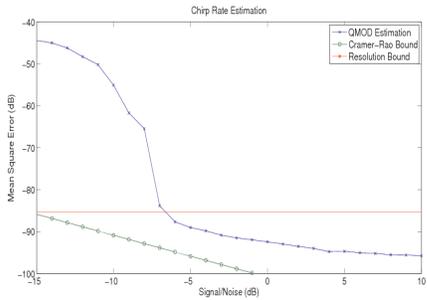
Peak-to-Parameter Mapping Region: Application of the DFRFT to chirp parameter estimation is not meaningful if a complete analysis of the invertibility of mapping is ignored. Therefore we looked at the peak-to-parameter mapping region in the $\alpha - \omega$ plane to see where the mappings satisfied the connectivity and adjacency conditions. The connectivity criteria is satisfied when the set of all chirp parameters that map to a single location in the chirp-rate versus center-frequency plane form a connected set, and adjacency criteria is satisfied when locations which are adjacent in the transform plane map to adjacent regions in the chirp parameter space. We calculated the percentage of mapping pixels, where connectivity and adjacency conditions were satisfied. This will enable determination of the method resulting in the best valid mapping region in comparison to the expected mapping region shown in figure 8. Figure 9 depicts the percentage of pixels in $\alpha - \omega$ plane where both connectivity and adjacency criteria are fulfilled.

From these mapping regions, we can see that the connectivity and adjacency conditions are not fulfilled for large regions in the case of the bilinear transformation method, the S-matrix method and the higher order S-matrix method. The regions in which these two conditions are violated overlap in these three cases. In the case of the Grunbaum method, we can see that, the two conditions are not satisfied in two different regions. The connectivity criteria is almost satisfied in the diamond region for the Grunbaum method, but the adjacency criteria is not satisfied as expected in the diamond region. Finally, both the criteria are satisfied in almost entire region of the $\alpha - \omega$ plane for the case of the QMOD method.

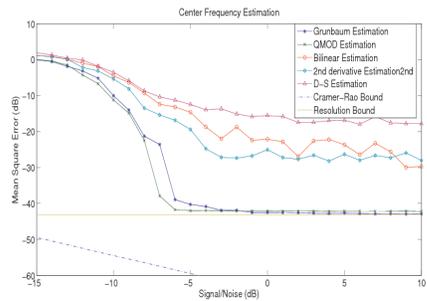
From this comparison, we can observe that the QMOD technique is the only method where the percentage of mapping pixels satisfying both the connectivity and adjacency criteria, increases as N increases. We saw in section 3 that the linearity of the eigenvalue spectrum increases when N increases only in the case of the QMOD basis. We also observed that the percentage of eigenvalue linearity for the QMOD method is far above than that of all other methods. From figure 9, we can see that the percentage of pixels reaches almost 90% for the QMOD method for $N = 512$, which is far better



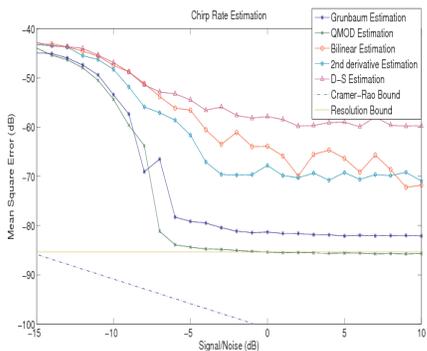
(a)



(b)

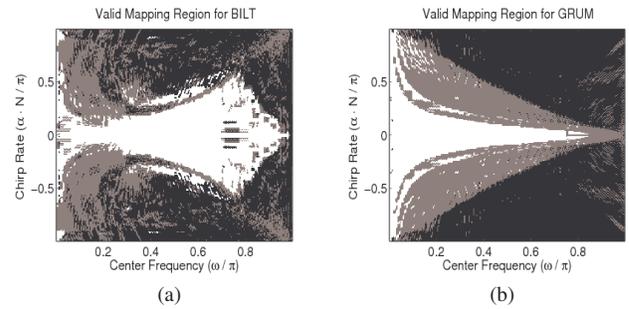


(c)

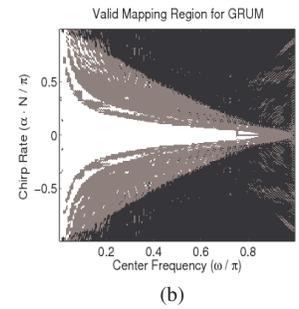


(d)

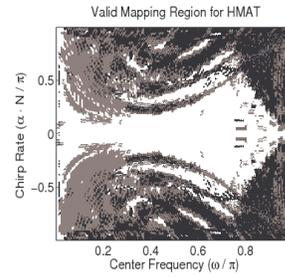
Fig. 7: MSE performance: (a,b):Cross-hair estimation error, (c,d):2D parameter estimation error. The MSE was calculated at each SNR using 1000 chirps of length $N = 256$, in the 'safe' range of $|\alpha|(N - 1) + |w| = IF < 0.85\pi$. A transform of size $N \times N$ was used, refined using minimum norm subspace decomposition and FFT of size $R = 4096$.



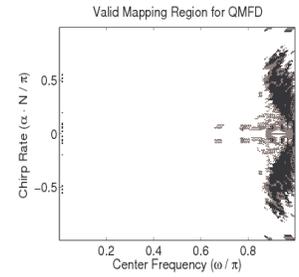
(a)



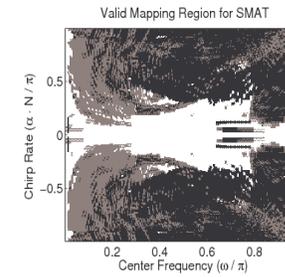
(b)



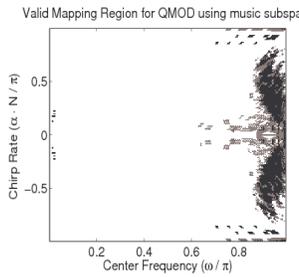
(c)



(d)



(e)



(f)

Fig. 8: Valid Mapping Regions for $N = 256$ obtained from (a) Bilinear Transformation Method, (b) Grunbaum Method, (c) Higher order S-matrix Method, (d) QMOD Method min-norm and (e) Dickinson-Steiglitz Method (f) QMOD Method using MUSIC.

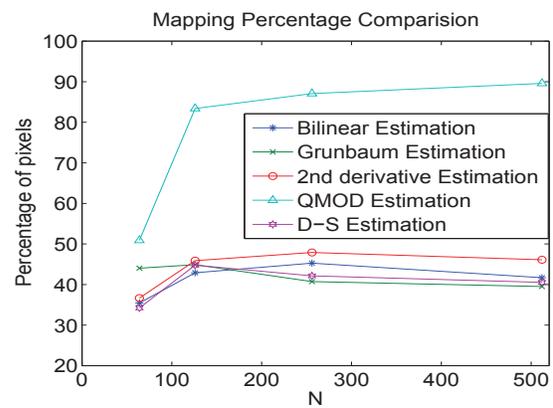


Fig. 9: Plot of percentage of pixels in $\alpha - \omega$ plane where both connectivity and adjacency criteria are fulfilled. The QMOD method exhibits a steady increase in invertibility of the peak-to-parameter mapping with increase in N .

than all the other methods. This confirms the fact that the QMOD basis produces the least estimation error because it has the largest valid peak-to-parameter mapping region in the $\alpha - \omega$ plane. We can also say that the valid mapping region will cover the entire $\alpha - \omega$ plane if N is sufficiently large for the QMOD basis.

4. CONCLUSIONS

Among the five commuting matrix methods for furnishing a basis of DFT eigenvectors, needed for computing the DFRFT, discussed in this paper, the QMOD method, a method in which the diagonal matrix \mathbf{Q} takes its diagonal as the zero crossings of the N th order G-H function instead of taking its diagonal as equally spaced values used in the conventional QMFD method, produces the sharpest peak for single chirp application. This result was verified by using different metrics such as mainlobe-to-sidelobe ratio, the 10-dB bandwidth, and quality factor. Also, the QMOD method produces the most linear eigenvalue spectrum. The same method results in less error-norm than the other methods, which confirms that the eigenvectors of the QMOD matrix are closest to that of the continuous G-H operator.

As is evident from the previous section, the QMOD method produces the least parameter estimation error for both center frequency estimation and chirp rate estimation. This estimation error is the closest to the Cramer-Rao lower bound when it is calculated for the QMOD method in combination with the cross-hair estimation technique. As we saw in the previous section, the QMOD method has an invertibility region of almost 90% of the $\alpha - \omega$ plane, which is far better than those of the other four methods. Therefore, we can conclude that the deviation from a fully linear eigenvalue spectrum of the DFT commuting matrix produces a large proportion of peak-to-parameter mapping pixels where the invertibility criteria are violated, and loss of invertibility results in larger chirp parameter estimation errors.

5. REFERENCES

- [1] W. H. S.C. Pei and J. Ding, "Ft-commuting matrix with arbitrary or infinite order second derivative approximation," *IEEE Trans. on Sig. Proc.*, vol. 51, no. 1, January 2009.
- [2] A. Serbes and L. Dural-Ata, "Eigenvectors of the discrete fourier transform based on the bilinear transform," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, no. 191085, June 2010.
- [3] B. Santhanam and T. Santhanam, "On discrete gauss-hermite functions and eigenvectors of the discrete fourier transform," *Signal Processing, Elsevier Science*, vol. 88, no. 6, pp. 2738–2746, November 2008.
- [4] D. Peacock and B. Santhanam, "Comparison of centered discrete fractional fourier transform for chirp parameter estimation," *Proc. of IEEE DSP/SPED Workshop*, August 2013.
- [5] B. Dickinson and K. Steiglitz, "Eigenvectors and functions of the discrete fourier transform," vol. 30, February 1982.
- [6] J. Vargas-Rubio and B. Santhanam, "On the grunbaum commutator based discrete fractional fourier transform," *Proc. of ICASSP-04*, vol. 2, pp. 641–644, May 2004.
- [7] F. Grunbaum, "The eigenvectors of the discrete fourier transform: A version of the hermite functions," *Journal of Math. Anal. Appl.*, vol. 88, no. 2, pp. 355–363, August 1982.
- [8] S. Clary and D. Mugler, "Shifted fourier matrices and their tri-diagonal commutators," *SIAM Journal Matrix Analysis and Applications*, vol. 24, no. 3, pp. 809–821, January 2003.
- [9] B. Santhanam and T. Santhanam, "Discrete gauss hermite functions and eigenvectors of the centered discrete fourier transform," *Proc. of ICASSP*, pp. 1385–1388, April 2007.
- [10] B. Santhanam and M. Hayat, "On a pseudo-subspace framework for discrete fractional Fourier transform based chirp parameter estimation", *Processing of IEEE DSP/SPE Workshops*, pp. 360-363, January 2011.
- [11] D. Peacock and B. Santhanam, "Multicomponent subspace chirp parameter estimation using discrete fractional fourier analysis," *Processing of IASTED Conference on SIP-2011*, pp. 326–333, December 2011.