Constraints to solve parallelogram grid problems in 2D non separable linear canonical transform

Liang Zhao[a,*], John J Healy[b], Inbarasan Muniraj[b], Xiao-Guang Cui[c], Ra'ed Malallah[b,d], James P Ryle[b], and John T. Sheridan[b].

[a]The Insight Centre for Data Analytics, University College Dublin, Belfield, Dublin 4, Ireland.

[b]School of Electrical and Electronic Engineering, IoE[2] Lab, SFI-Strategic Research Cluster in Solar Energy Conversion, College of Engineering and Architecture, University College Dublin, Belfield, Dublin 4, Ireland.

[c]Institute of Automation, Chinese Academy of Sciences, No. 95, East Zhong Guan Cun Road, Hai-Dian District, Beijing, 100190, P.R.China.

[d]Physics Department, Faculty of Science, University of Basrah, Garmat Ali, Basrah, Iraq.

Corresponding author: liang.zhao@ucd.ie

ABSTRACT

The 2D non-separable linear canonical transform (2D-NS-LCT) can model a range of various paraxial optical systems. Digital algorithms to evaluate the 2D-NS-LCTs are important in modeling the light field propagations and also of interest in many digital signal processing applications. In [Zhao 14] we have reported that a given 2D input image with rectangular shape/boundary, in general, results in a parallelogram output sampling grid (generally in an affine coordinates rather than in a Cartesian coordinates) thus limiting the further calculations, e.g. inverse transform. One possible solution is to use the interpolation techniques; however, it reduces the speed and accuracy of the numerical approximations. To alleviate this problem, in this paper, some constraints are derived under which the output samples are located in the Cartesian coordinates. Therefore, no interpolation operation is required and thus the calculation error can be significantly eliminated.

Keywords: Digital holography, ABCD transforms, Numerical approximation and analysis, Discrete optical signal processing.

1. INTRODUCTION

The continuous two dimensional *non-separable* linear canonical transform (2D-NS-LCT) of an input signal $g(x,y)$ can be given by [1, 2],

$$\begin{aligned} \mathrm{G}(x',y') &= L_M\{g(x,y)\}(x',y') \\ &= \frac{1}{\sqrt{j\det(B)}}\iint_{-\infty}^{\infty}\exp\left[\frac{j\pi\left(k_1x'^2+k_2x'y'+k_3y'^2\right)}{\det(B)}\right]\exp\left\{\frac{j2\pi[(-b_{22}x'+b_{12}y')x+(b_{21}x'-b_{11}y')y]}{\det(B)}\right\} \\ &\quad\times\exp\left[\frac{j\pi\left(p_1x^2+p_2xy+p_3y^2\right)}{\det(B)}\right]g(x,y)dxdy, \end{aligned} \tag{1-Ia}$$

where M is the transformation matrix of the LCT system,

$$M=\left(\begin{array}{cc:cc} a_{11} & a_{12} & b_{11} & b_{12} \\ a_{21} & a_{22} & b_{21} & b_{22} \\ \hdashline c_{11} & c_{12} & d_{11} & d_{12} \\ c_{21} & c_{22} & d_{21} & d_{22} \end{array}\right)=\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{1-Ib}$$

where A, B, C and D are 2×2 sub-matrices, e.g. $B=\{b_{11}b_{12};b_{21}b_{22}\}$. In this case it is assumed that $\det(B)\neq 0$, where $\det(B)$ represents the determinant of sub-matrix B. $k_1\sim k_3$ and $p_1\sim p_3$ are as defined in [1].

Holography: Advances and Modern Trends V, edited by Miroslav Hrabovský, John T. Sheridan, Antonio Fimia, Proc. of SPIE Vol. 10233, 102331V · © 2017 SPIE · CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2265862

$$k_1 = d_{11}b_{22} - d_{12}b_{21},\ k_2 = 2(-d_{11}b_{12} + d_{12}b_{11}), k_3 = -d_{21}b_{12} + d_{22}b_{11},$$
$$p_1 = a_{11}b_{22} - a_{21}b_{12},\ p_2 = 2(a_{12}b_{22} - a_{22}b_{12}),\ p_3 = -a_{12}b_{21} + a_{22}b_{11}, \quad \text{(1-Ic)}$$

Definition of the 2D-NS-LCT when $B = 0$,$\Longrightarrow \det(B) = 0$, is given by [3],

$$\begin{aligned} \mathrm{G}(x', y') &= L_M\{g(x,y)\}(x', y') \\ &= \sqrt{\det(D)}\exp\left[j\pi(c_{11}d_{11} + c_{12}d_{12})x'^2\right]\exp[j2\pi(c_{11}d_{21} + c_{12}d_{22})x'y'] \\ &\quad \times \exp[j\pi(c_{21}d_{21} + c_{22}d_{22})y'^2]\, g(d_{11}x' + d_{21}y', d_{12}x' + d_{22}y'), \end{aligned} \quad \text{(1-Id)}$$

where $M = \{D^{T^{-1}}\ \ 0; C\ D\}$. Definition of the 2D-NS-LCT when $\det(B) = 0$ but $B \neq 0$ are given in [3].

The inverse 2D-NS-LCT recovers $g(x, y)$ from $G(x', y')$ by,

$$g(x, y) = L_{M^{-1}}\{G(x', y')\}(x, y), \quad \text{(1-II)}$$

where $M^{-1} = \{D^T\ -B^T;\ -C^T\ A^T\}$, and D^T is the transpose of sub-matrix D.

The continuous NS-LCT is *additive* [1, 2]:

$$\mathrm{L}_{\mathrm{M}_2}\left\{\mathrm{L}_{\mathrm{M}_1}\{\mathrm{g}\}\right\} = \mathrm{L}_{\mathrm{M}_3}\{\mathrm{g}\}, \quad \text{(1-III)}$$

where $M_3 = M_2M_1$. When $M_2 = M_1^{-1}$, we obtain a statement that the transform is *unitary* [1, 2]:

$$L_{M^{-1}}\{L_M\{g\}\} = g. \quad \text{(1-IV)}$$

A wide variety of 2D orthogonal paraxial optical systems can be modelled by the 2D separable linear canonical transform (2D-S-LCT) [2]. For example, the Fourier transform, the Fresnel transform, the fractional Fourier transform are its special cases. However many more general paraxial optical systems exist which cannot be modelled using the 2D-S-LCT [2]. For example, systems which:

(i) are non-orthogonal, non-axially symmetric, or contains an anamorphic lens [4, 5];
(ii) Involve coupling and/or shearing operations (between components along the different dimensions), e.g., a coordinate transform, see Table II in [1]; and
(iii) Involve rotations between any arbitrary planes in phase space, e.g., the gyrator transform [6~10];

Such systems can only be modelled using the 2D-NS-LCT [2]. They may be distinguished by their ABCD matrices. A separable system has diagonal A, B, C, and D sub-matrices, and hence zero-valued elements in the following positions in its ABCD matrix:

$$\begin{pmatrix} a_x & 0 & b_x & 0 \\ 0 & a_y & 0 & b_y \\ c_x & 0 & d_x & 0 \\ 0 & c_y & 0 & d_y \end{pmatrix} \quad \text{(1-V)}$$

We then associate the matrix $\begin{pmatrix} a_x & b_x \\ c_x & d_x \end{pmatrix}$ with the x-direction and the matrix $\begin{pmatrix} a_y & b_y \\ c_y & d_y \end{pmatrix}$ with the y-direction. Non-separable systems are fundamentally different and have non-zero values in one or more of these off-diagonal matrix positions. These values act to create a link (or coupling) between what happens along the two directions.

According to the additive property, see Eq. (1-III), the ABCD matrix can be decomposed in many ways. One such decomposition is reproduced here, see [20, 21],

$$M = M_4M_3M_2M_1 = \begin{pmatrix} B & 0 \\ 0 & B^{T^{-1}} \end{pmatrix}\begin{pmatrix} I & 0 \\ B^{\mathrm{T}}D & I \end{pmatrix}\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\begin{pmatrix} I & 0 \\ B^{-1}A & I \end{pmatrix}. \quad \text{(1-VI)}$$

Consider the effect of this series of operations on a discrete input signal. The chirp multiplication (M_1) does not alter the nature of the signal, i.e. it remains a discrete signal, and the locations of the samples are unchanged. The FT (M_2) can be efficiently performed using an FFT, and hence the output of this operation is also discrete. The 2nd chirp multiplication (M_3) does not change the location of the signal samples. The samples are now located on a regular Cartesian grid, much as at the start of the calculation. However, the coordinate transform (M_4) changes the locations of these samples to potentially "inconvenient" or "irregular" locations. This creates the problems that this paper attempts to address.

Except in rare occasions, e.g., when B is a diagonal or anti-diagonal matrix, the output image shape of the corresponding NS optical system is skewed, i.e. a rectangular input image becomes a parallelogram. This presents difficulties in a number of situations. For example:

(1) It is common to assume that the input to a discrete transform has been sampled on a regular grid. The effect of the coordinate transform (M_4) is to move these samples to a skewed rectilinear grid. As a result, the output of the transform cannot be used as an input to another transform without further preprocessing. Interpolation techniques can be used to regularize the sample grid while this increases the number of samples, requires computational resources and may not be perfectly reversible. As a consequent, the output of the transform is not immediately suitable to act as the input for further calculation – it is literally a parallelogram peg for a square hole. This is relevant to calculations involving several stages or iterative algorithms where the calculation process involves going repeatedly back and forth between sample values in transform related domains. This problems must be solved before one can deal with considerations of unitarity or additivity considered elsewhere in this paper; and
(2) Similarly, this coordinate transform occurs when one wishes to display the resulting output on a digital screen (made up of regularly arranged pixels). Once again interpolation becomes necessary to fit the data to the display.

Note that this problem does not occur in both the 1D transform case and 2D separable transform case because the coordinate transform reduces to a simple magnification, which is easily accounted for in terms of sampling rates.

The effect of the coordinate transform on the locations of the samples is further illustrated in Fig. 1. For the input shown in Fig. 1(a1) the effects of three different transforms are presented. The outputs appear in Figs. 1(b1) ~ (d1), and the associated sampling grids in Figs. 1(b2) ~ (d2). The three transforms are as follows:

Fig. (b) The simplest case. A separable system or a system where B is a diagonal or an anti-diagonal matrix;
Fig. (c) The most common case among non-separable systems. The difficulty discussed above can be clearly seen;
Fig. (d) An example of the special non-separable case that has the advantageous sample arrangement under discussion.

The remainder of this paper is structured as follows. In Section 2 a constraint on the ABCD parameters and input signal is proposed which guarantees alike results in Figs. 1(b1) and 1(b2) is output. In other words, while the output signal is not rectangular, it does consist of samples located on a regular grid. Such a data set can be easily displayed on a typical discrete display such a liquid-crystal display screen. Finally, in Section 3, we conclude our discussions.

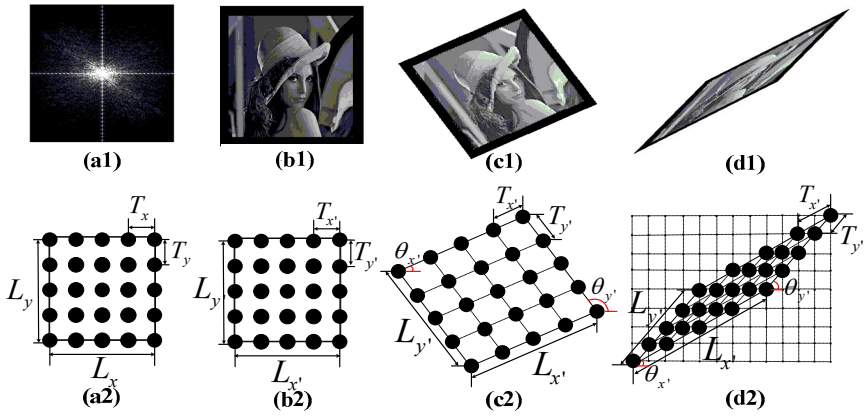

Fig. 1. Illustration of sampling intervals in 2D-NS-LCT: (a1) the absolute value of an example (frequency) input image. It is sampled on a regular grid, like that of (a2). The remaining images

[(b1), (c1) and (d1)] show the absolute value of several different transform of the data. In (b1), the transform is an S-LCT (or a NS-LCT system where B is a diagonal or anti-diagonal matrix), and hence the corresponding output sampling grid is also a regular grid, see (b2). In (c1), the transform is a typical NS-LCT. Its samples are located on a rectilinear but non-Cartesian grid as shown in (c2). Finally, the transform in (d1) is a special NS-LCT case, where the skewed samples coincide with a Cartesian grid, as depicted in (d2).

2. PROPOSED CONSTRAINT

In this section, we propose a constraint under which the output samples can be located in a regular Cartesian coordinate system. For ease of presentation, the detailed derivation of the constraint is given in detail in Section 3. At the start of this section the constraint is simply stated.

The constraint: The output samples of a 2D-NS-LCT can lie on a Cartesian coordinates without interpolation, if any of the seven constraints in Table I is satisfied:

TABLE I The proposed constraint

	Constraints
①B is a diagonal matrix	$b_{12} = b_{21} = 0$
②B is an anti-diagonal matrix	$b_{11} = b_{22} = 0$
③B is an lower diagonal matrix	$b_{12} = 0, \quad L_y = lb_{22}L_x/b_{21}$, and $l \in \mathbb{N}$
④B is a upper diagonal matrix	$b_{21} = 0, \quad L_y = b_{12}L_x/(lb_{11})$, and $l \in \mathbb{N}$
⑤$b_{11} = 0$	$L_y = \lvert b_{22}L_x/b_{21} \rvert$
⑥$b_{22} = 0$	$L_y = lb_{12}L_x/b_{11}, \quad l \in \mathbb{N}$
⑦ $b_{ij} \neq 0\ \forall\ i,j$	$\frac{b_{11}b_{22}}{b_{12}b_{21}} = 1 + \frac{1}{kl}, L_y = l\frac{b_{22}}{b_{21}}L_x$, and $k, l \in \mathbb{N}$

Recall that in relation to Table I, $b_{11}, b_{12}, b_{21}, b_{22}$ are the four elements of sub-matrix B, which is assumed to have non-zero determinant. It should be noted that in Table I:

- In Case ① or ②, when the matrix B is a diagonal or anti-diagonal matrix, the resulting output samples lie on regular grid without further constraint. The only difference between the input and output grids is that the sides are in general of different length;
- In cases ③ $b_{12} = 0$, and ④ $b_{21} = 0$, the matrix B is a lower or upper diagonal matrix. In each case a further constraint must be satisfied to make the output samples lie on a regular grid, i.e. ③ $L_y = lb_{22}L_x/b_{21}$, or ④ $L_y = \frac{b_{12}L_x}{lb_{11}}$, where $l \in \mathbb{N}$, and L_x and L_y represent the extents of the input signal along the x and y directions;
- Similarly, in cases ⑤ $b_{11} = 0$, and ⑥ $b_{22} = 0$, a further constraint should be satisfied to get a regular output grid: ⑤ $L_y = |b_{22}L_x/b_{21}|$, or ⑥ $L_y = lb_{12}L_x/b_{11}$, where $l \in \mathbb{N}$; and
- In case ⑦ $b_{ij} \neq 0\ \forall\ i,j$, i. e. $b_{12} \neq 0, b_{21} \neq 0, b_{22} \neq 0$, the additional constraints $\frac{b_{11}b_{22}}{b_{12}b_{21}} = 1 + \frac{1}{kl}$ and $L_y = \frac{lb_{22}L_x}{b_{21}}$, where $k, l \in \mathbb{N}$, should be satisfied to get a regular output grid.

3. PROPOSED CONSTRAINT DERIVATION

In this section, the proposed constraints in Table 1 are derived.

3.1 The effect of shearing operation

First, let us begin by discussing the effect of shearing operation on output sampling grid. Given the following shearing operation:

$$\begin{pmatrix} 1 + shxshy & shx \\ shy & 1 \end{pmatrix} = \begin{pmatrix} 1 & shx \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ shy & 1 \end{pmatrix}, \tag{3-I}$$

i.e. a shearing on y- axis followed by a shearing on x- axis, see Eq. (3-I) from right to left sides. In this

paper we assume that the number of input samples on x- and y- axis are the same, denoted as N.

In the case when the input sampling intervals $T_x = T_y$, i.e. the input window is square, the Cartesian output grid can be obtained if in Eq. (3-I)

$$shx \text{ and } shy \text{ are integers.} \tag{3-II}$$

Further in the case when $T_y = pT_x$, i.e. the input window is rectangle, where p is positive real number, the Cartesian output grid can be acheived if in Eq. (3-I)

$$shx = k/p, shy = lp, \text{and } k, l \in \mathbb{N} \tag{3-III}$$

3.2 The effect of coordinate transform (of the 2D-NS-LCT)

Based on the discussions in subsection 3.1, let us now consider the coordinate transform B of the 2D-NS-LCT. In [1] we have shown that when $b_{22} \neq 0$ and $\det(B) \neq 0$, the sub-matrix B can be decomposed into,

$$\begin{pmatrix} b_{11} b_{12} \\ b_{21} b_{22} \end{pmatrix} = \begin{pmatrix} 1 + \frac{b_{12} b_{21}}{\det(B)} & \frac{b_{12}}{b_{22}} \\ \frac{b_{21} b_{22}}{\det(B)} & 1 \end{pmatrix} \begin{pmatrix} \frac{\det(B)}{b_{22}} & 0 \\ 0 & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & \frac{b_{12}}{b_{22}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{b_{21} b_{22}}{\det(B)} & 1 \end{pmatrix} \begin{pmatrix} \frac{\det(B)}{b_{22}} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & b_{22} \end{pmatrix}. \tag{3-IV}$$

Eq. (3-IV) indicates that B can be performed by taking a scaling on y; scaling on x; shearing on y; then shearing on x, see Eq. (3-IV) from the right to left sides. We note that:

(1) The two scaling operations have no critical effect on the output sampling grid shape, i.e. the output grid after these two scaling operations is still rectangle, only its side lengths values are changed. The corresponding output rectangle grid area is $\frac{|\det(B)|}{T_x T_y}$, where $|\det(B)|$ returns the absolute value of the determinant of B.
We note that the input grid area value is $T_x T_y N^2$. After the FT operation, see Eq. (1-VI), the output grid area is $\frac{1}{T_x T_y}$. Then after the two scaling operations in Eq. (3-IV) the corresponding output grid area value is $\frac{|\det(B)|}{T_x T_y}$.

(2) Therefore, it is evident that two shearing operations play an important role in the output shape. However we note that the shearing operation has no effect on the output image area, i.e. the area value of the final output parallelogram is still $\frac{|\det(B)|}{T_x T_y}$.

Let us assume that $T_y = pT_x$, i. e. $L_y = pL_x$. From Eq. (3-IV), after the FT and two scaling operations (in Eq. (3-IV)), the side lengths of the new output sampling grid $L_{x'}$ and $L_{y'}$ can be given by,

$$L_{x'} = \frac{\det(B)}{b_{22} T_x}, \text{and } L_{y'} = \frac{b_{22}}{T_y} = \frac{b_{22}}{pT_x}. \tag{3-V}$$

i.e.

$$L_{y'} = p' L_{x'} \tag{3-VI}$$

where $p' = \frac{b_{22}^2}{p\det(B)}$. According to Eqs. (3-III), (3-IV), and (3-VI), the Cartesian output grid can be obtained if

$$\frac{b_{12}}{b_{22}} = \frac{k}{p'} = k \frac{p\det(B)}{b_{22}^2}, \ \frac{b_{21} b_{22}}{\det(B)} = lp' = l \frac{b_{22}^2}{p\det(B)}, \text{and } k, l \in \mathbb{N}. \tag{3-VII}$$

The constraints for the following 5 cases in Table 1 can be obtained from Eq. (3-VII).

① In **the diagonal case**, $\begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & b_{22} \end{pmatrix}$, i.e. only scaling factors on x- and y- axis are involved here, the resulting output samples always lie on Cartesian grid;

③ In **the lower triangular case** ($b_{12} = 0$), Eq. (3-VII) can be simplified as $\frac{b_{21} b_{22}}{\det(B)} = l \frac{b_{22}^2}{p\det(B)}$, i.e. $p = \frac{l b_{22}}{b_{21}}$. Therefore, the constraint is $L_y = pL_x = l \frac{b_{22}}{b_{21}} L_x$, where $l \in \mathbb{N}, (k = 0)$;

④ In **the upper triangular case** ($b_{21}=0$), Eq. (3-VII) can be simplified as $\frac{b_{12}}{b_{22}}=k\frac{p\det(B)}{b_{22}^2}$, i.e. $p=\frac{b_{12}}{kb_{11}}$. Therefore, $L_y=pL_x=\frac{1}{k}\frac{b_{12}}{b_{11}}L_x$, where $k\in\mathbb{N},(l=0)$;

⑤ In **the anti lower triangular case** ($b_{11}=0$), Eq. (3-VII) can be simplifed as $p=\frac{b_{22}}{-kb_{21}}=\frac{lb_{22}}{-b_{21}}$, i.e. $p=\left|\frac{b_{22}}{b_{21}}\right|$. Therefore, $L_y=pL_x=\left|\frac{b_{22}}{b_{21}}\right|L_x$.

⑦ In **the case when $b_{ij}\neq 0,\forall i,j$,** (i.e. in Eq. (3-IV): $\frac{b_{12}}{b_{22}}\neq 0$ and $\frac{b_{21}b_{22}}{\det(B)}\neq 0$), Eq. (3-VII) can be rewritten as

$$\det(B)=\frac{b_{12}b_{21}}{kl},\text{ and } p=\frac{lb_{22}}{b_{21}},\text{ and } k,l\in\mathbb{N}, \tag{3-IX}$$

i.e. $\frac{b_{11}b_{22}}{b_{12}b_{21}}=1+\frac{1}{kl}$, $L_y=pL_x=l\frac{b_{22}}{b_{21}}L_x$.

In the 2nd and the 6th cases in Table 1, the decomposition of the matrix B given in Eq. (3-IV) is invalid because of $b_{22}=0$.

② In **the anti-diagonal case**, the matrix B can be decomposed into:

$$\begin{pmatrix}0 & b_{12}\\ b_{21} & 0\end{pmatrix}=\begin{pmatrix}b_{12} & 0\\ 0 & b_{21}\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}=\begin{pmatrix}b_{12} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & b_{22}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \tag{3-X}$$

i.e. a 90° rotation transform followed by a reflect about x-axis, scaling on y, and then scaling on x. It can be seen that no shearing operation is involved here, therefore the Cartesian output grid can always been obtained in this case.

⑥ In **the anti upper triangular case** ($b_{22}=0$),

$$\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & 0\end{pmatrix}=\begin{pmatrix}0 & b_{12}\\ b_{21} & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}=\begin{pmatrix}1 & \frac{b_{11}}{b_{21}}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}b_{21} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & b_{12}\end{pmatrix}, \tag{3-XI}$$

i.e. a scaling on y, scaling on x, 90° rotation, reflect about x-axis, and shearing on x. Only the final shearing operation determines whether the final output samples can still lie on Cartesian grid. The output extents $L_{x'}$ and $L_{y'}$ before the shearing operation in Eq. (3-XI), (i.e. the resulting output extents after the FT in Eq. (1-VI) and the former four operations in Eq. (3-XI)), can be given by,

$$L_{x'}=\frac{b_{12}}{T_y},\text{ and } L_{y'}=\frac{b_{21}}{T_x}, \tag{3-XII}$$

i.e. $L_{y'}=p'L_{x'}$, where $p'=\frac{L_{y'}}{L_{x'}}=\frac{b_{21}T_y}{b_{12}T_x}=\frac{b_{21}}{b_{12}}p$. According to Eq. (3-III), the Cartesian output grid can be obtained if the shearing operation parameter $\frac{b_{11}}{b_{21}}=\frac{k}{p'}=\frac{kb_{12}}{b_{21}p}$, i.e. $p=k\frac{b_{12}}{b_{11}}$. Therefore the constraint for this case is $L_y=pL_x=k\frac{b_{12}}{b_{11}}L_x$, where $k\in\mathbb{N}$.

4. FUTURE WORK

In this paper, we have introduced and derived the constraints that result in unitary discrete 2D-NS-LCT algorithm. Based on the proposed constraint in this paper, a fast unitary algorithm has been proposed in [20]. In [20] we have demonstrated that the unitary 2D-NS-LCT significantly improves the performance of the iterative phase retrieval algorithm.

The proposed constraints in this paper are derived by means of basic geometry. The question is have we found all unitary discrete 2D-NS-LCT? Our future work will focus on exploring the sufficient constraints for the discrete LCTs to ensure unitarity.

REFERENCES

[1] L. Zhao, J. J. Healy, and J. T. Sheridan, "The 2D non-separable linear canonical transform: Sampling theorem and unitary discretization," J. Opt. Soc. Am. A **31**, 2631-2641 (2014).

[2] A. Koç, H. M. Ozaktas, and L. Hesselink, "Fast and accurate computation of two-dimensional non-separable quadratic-phase integrals," J. Opt. Soc. Am. A **27**, 1288-1302 (2010).
[3] S.-C. Pei, "Two-dimensional affine generalized fractional Fourier transform," IEEE Trans. Sig. Proc. **49**, 878-897 (2001).
[4] A. E. Sieg, *Lasers*, University Science Books, Palo Alto, 1986.
[5] G. Kloos, [Matrix methods for optical layout], SPIE, Bellingham, 59-75, 2007.
[6] J. A. Rodrigo, T. Alieva, and M. L. Calvo, "Optical system design for orthosymplectic transformations in phase space," J. Opt. Soc. Am. A **23** (10), 2494-2500 (2006).
[7] J. A. Rodrigo, T. Alieva, and M. L. Calvo, "Experimental implementation of the gyrator transform," J. Opt. Soc. Am. A **24**, 3135–3139 (2007).
[8] J. A. Rodrigo, T. Alieva, and M. L. Calvo, "Gyrator transform: Properties and applications," Opt. Express **15**, 2190–2203 (2007).

[9] J. A. Rodrigo, T. Alieva, and M. L. Calvo, "Applications of gyrator transform for image processing," Opt. Commun. **278**, 279–284 (2007).
[10] S.-C. Pei, and J.-J. Ding, "Properties, digital implementation, applications, and self-image phenomena of the gyrator transform," 17th European Signal processing conference, 441-445 (2009).
[11] J.-J. Ding, S.-C. Pei, and C.-L. Liu, "Improved implementation algorithms of the two-dimensional non-separable linear canonical transform," J. Opt. Soc. Am. A **29**, 1615-1624 (2012).
[12] A. V. Oppenheim, and R. W. Schafer, Discrete-time signal processing, 2nd-ed., Prentice Hall, New Jersey, 1999.
[13] L. Zhao, J. J. Healy, and J. T. Sheridan, "Unitary discrete linear canonical transform: Analysis and application," Appl. Opt.**52** (7), C30-C36 (2013).
[14] J. R. Fienup, "Phase retrieval algorithms: A comparison," Applied Opt. **21** (15), 2758-2769 (1982).
[15] N. M. Atakishiyev, G. S. Pogosyan, L. E. Vicent, and K. B. Wolf, "Finite two-dimensional oscillator: I. the Cartesian model," J. Phys. A: Math. Gen. **34** (44), 9381-9398 (2001).
[16] K. B. Wolf, and T. Alieva, "Rotation and gyration of finite two-dimensional modes," Opt. Soc. Am. A**25** (2), 365-370(2008)
[17] L. E. Vicent, and K. B. Wolf, "Analysis of digital images into energy-angular momentum modes," J. Opt. Soc. Am. A**28** (5), 808-814 (2011).
[18] S.-C. Pei, "Discrete spherical harmonic oscillator transforms on the Cartesian grids using transformation coefficients," IEEE Trans. Sig. Proc. **61** (5), 1149-1164 (2013).
[19] G. Krötzsch, K. Uriostegui, and K. B. Wolf, "Unitary rotations in two-, three-, and D-dimensional Cartesian data arrays," J. Opt. Soc. Am. A. **31** (7), 1531-1535 (2014).
[20] L. Zhao, J. T. Sheridan, and J. J. Healy, "Unitary algorithm for non-separable linear canonical transforms applied to iterative phase retrieval," accepted by IEEE Signal Proc. Let. (2017).
[21] J. J. Healy, M. A. Kutay, H. M. Ozaktas, and J. T. Sheridan, [Linear Canonical Transforms], Springer, New York (2016).