applied optics

Effect of spatial distribution of partial information on the accurate recovery of optical wave fields

Figen S. Oktem[1,2,*] and Haldun M. Ozaktas[2]

[1]Department of Electrical and Electronics Engineering, Middle East Technical University (METU), TR-06800 Cankaya, Ankara, Turkey
[2]Department of Electrical Engineering, Bilkent University, TR-06800 Bilkent, Ankara, Turkey
*Corresponding author: figeno@metu.edu.tr

Received 6 September 2016; revised 10 November 2016; accepted 14 November 2016; posted 15 November 2016 (Doc. ID 275197); published 19 December 2016

We consider the problem of recovering a signal from partial and redundant information distributed over two fractional Fourier domains. This corresponds to recovering a wave field from two planes perpendicular to the direction of propagation in a quadratic-phase multilens system. The distribution of the known information over the two planes has a significant effect on our ability to accurately recover the field. We observe that distributing the known samples more equally between the two planes, or increasing the distance between the planes in free space, generally makes the recovery more difficult. Spreading the known information uniformly over the planes, or acquiring additional samples to compensate for the redundant information, helps to improve the accuracy of the recovery. These results shed light onto redundancy and information relations among the given data for a broad class of systems of practical interest, and provide a deeper insight into the underlying mathematical problem.

OCIS codes: (070.0070) Fourier optics and signal processing; (070.2575) Fractional Fourier transforms; (100.3190) Inverse problems.

https://doi.org/10.1364/AO.56.00A133

1. INTRODUCTION

As is well known, if we know the values of a propagating wave field on a planar surface, we can determine the values of the wave field over the entire three-dimensional space. The only requirement is that there are a sufficient number of known sample values within its extent to satisfy sampling constraints. Here we consider the problem where known sample values do not all lie on a single plane but may be distributed over more than one plane. In this case, the known sample values may not represent independent pieces of information. In particular, a number of samples that would have been sufficient had they all been lying on a single plane perpendicular to the direction of propagation, may now not be sufficient to uniquely determine the wave field. The purpose of this work is to develop insight into how the distribution of known sample points affects our ability to recover the whole field.

In order to bring forth the essence of the issue as transparently as possible, we will work with only one transverse dimension instead of two. Furthermore, rather than working with the Fresnel integral that describes the propagation of light under common assumptions, we will model propagation behavior with the fractional Fourier transform (FRT) [1–3]. It has been well established that the Fresnel integral can be written as an FRT followed by simple scaling followed by a residual quadratic-phase factor, and the propagation of waves can be considered as a process of continual fractional Fourier transformation, where the fractional order monotonically increases as a function of the distance of propagation [4–6]. Planes perpendicular to the axis of propagation correspond to fractional Fourier domains (FRFDs) [3]. Furthermore, the same is true for all quadratic-phase systems [4], which are systems including an arbitrary sequence of lenses, sections of free space, and quadratic graded-index media [7–16]. Working with the FRT will allow us to approach the heart of the problem in its purest form and identify the main trends as clearly as possible, while the resulting important observations remain applicable to a broad class of systems of practical interest.

Thus we will consider a class of signal recovery problems where distributed partial information in some fractional Fourier domains are available and the aim is to recover the wave field everywhere [17–19]. Physically, this corresponds to the problem of recovering wave fields from partial information distributed over several transverse planes in either free space, or multilens quadratic-phase systems. These problems can arise in optical, acoustical, and electromagnetic wave propagation. One timely area where this problem is particularly relevant is digital three-dimensional television [20–26].

While some *ad hoc* algorithms have been applied to these problems, there has not been much theoretical progress.

A numerical solution to the problem in a pure fractional Fourier domain context has been given in [27], where an iterative algorithm has been developed based on the method of projections onto convex sets. A numerical approach to the recovery of the signal from its samples at arbitrarily distributed points has been presented in an optical context in [24]. An information theoretic approach to a related but differently formulated problem has been discussed in [28–31].

The purpose of this work is not to solve the resulting signal recovery problem numerically, but to develop insight into the nature of redundancy and information relationships in such inverse problems under different configurations of known information. We formulate the problem as a linear system of equations and by looking at the condition number of the system, try to reach general understanding of how the distribution of known information affects our ability to recover the wave field. The condition number of the system indicates how reliably such a system can be solved in the presence of inevitable modeling errors, sample errors, round-off errors, and finite machine precision. A poor condition number indicates there is not sufficient linearly independent information to reliably solve the problem.

Our analysis of conditioning also has implications for the choice of signal recovery methods. In particular, bad condition numbers indicate that taking a simple matrix inversion or least-squares approach may lead to inaccurate solutions. In this case, either regularization techniques that incorporate additional prior information into the solution should be exploited, or more samples should be taken, to improve conditioning [32–36].

In the following sections we first formulate the recovery problem in the language of linear algebra and then discuss how the condition number characterizes the difficulty of reliably solving the resulting system of equations. We then present the different observing scenarios we have investigated along with numerical results for them. Finally, we summarize our overall conclusions. Preliminary versions of this work appears in [17–19].

2. PROBLEM DEFINITION

The ath-order FRT of a function $f(x)$, denoted $f_a(x)$, is defined as [3]

$$f_a(x) \equiv (\mathcal{F}^a f)(x) \equiv \int_{-\infty}^{\infty} K_a(x,x') f(x') \mathrm{d}x', \tag{1}$$

$$K_a(x,x') \equiv A_\phi e^{i\pi(\cot\phi x^2 - 2\csc\phi x x' + \cot\phi x'^2)},$$

$$A_\phi = \sqrt{1 - i\cot\phi}, \qquad \phi = a\pi/2,$$

when $a \neq 2k$, and $K_a(x,x') = \delta(x-x')$ when $a = 4k$ and $K_a(x,x') = \delta(x+x')$ when $a = 4k \pm 2$, where k is an integer. The FRT is additive in index: $\mathcal{F}^{a_2}\mathcal{F}^{a_1} = \mathcal{F}^{a_2+a_1}$ and reduces to the ordinary Fourier transform and identity operators for $a = 1$ and $a = 0$, respectively. In the space-frequency plane, the ath-order FRT transforms a signal to the oblique axis making angle $\phi = a\pi/2$ with the space axis x. This axis is referred to as the ath-order FRFD [3]. As noted in the introduction, it has been shown that Fresnel propagation and all other quadratic-phase systems can be written as an FRT followed by simple scaling followed by a residual quadratic-phase factor, and the propagation of waves through such systems can be considered as a process of continual fractional Fourier transformation [4,8,14].

Consider two FRFDs of orders a_1 and a_2 such that the representation of a signal in these two domains is, respectively, denoted by $f(x)$ and $g(x)$. That is, $g(x) = (\mathcal{F}^a f)(x)$ with $a = a_2 - a_1$. We will assume that the signal f in the a_1th-order FRFD is approximately confined to the interval $[-\Delta x_{a_1}/2, \Delta x_{a_1}/2]$ and the signal g in the a_2th-order FRFD is approximately confined to the interval $[-\Delta x_{a_2}/2, \Delta x_{a_2}/2]$. That is, a sufficiently large percentage of the signal energy is confined to these intervals. Then, according to the FRT sampling theorem [37,38], the sampling interval in the a_1th-order domain should not be larger than $\delta x_{a_1} = 1/(\Delta x_{a_2} |\csc(a\pi/2)|)$. If we sample this domain at this rate, the total number of samples over the extent Δx_{a_1} will be given by $\Delta x_{a_1}/\delta x_{a_1} = \Delta x_{a_1}\Delta x_{a_2}|\csc(a\pi/2)|$. This product is a special case of the *bicanonical width product* [4,13]. Similarly, if we sample in the a_2th-order domain, the sampling interval should not be larger than $\delta x_{a_2} = 1/(\Delta x_{a_1} |\csc(a\pi/2)|)$. If we sample this domain at this rate, the total number of samples over the extent Δx_{a_2} will again be given by $\Delta x_{a_2}/\delta x_{a_2} = \Delta x_{a_1}\Delta x_{a_2}|\csc(a\pi/2)|$. Hence at each domain the signal can be sampled at N equally spaced points, where the number of samples N must be greater than or equal to $\Delta x_{a_1}\Delta x_{a_2}|\csc(a\pi/2)|$.

Let $\mathbf{f} = [f[-N/2], \ldots, f[N/2-1]]^{\mathrm{T}}$ and $\mathbf{g} = [g[-N/2], \ldots, g[N/2-1]]^{\mathrm{T}}$ denote the vectors of length N which represent the samples of the signals f and g in the a_1th- and a_2th-order FRFDs, respectively. If $a = a_2 - a_1$, the relation between the signals at these domains is given by

$$\mathbf{g} = \mathbf{F}^{\mathbf{a}}\mathbf{f}, \tag{2}$$

where $\mathbf{F}^a$ denotes the $N \times N$ ath-order discrete FRT matrix given in [39], cyclically shifted since we are working with discrete signals over $[-N/2, N/2-1]$ rather than $[0, N-1]$.

Let m_1 and m_2 denote the number of known samples in the a_1th- and a_2th-order FRFDs, respectively ($m_1 < N$ and $m_2 < N$). If the indices corresponding to the known samples form the vectors $\mathbf{k} = [k_1, \ldots, k_{m_1}]^{\mathrm{T}}$ and $\mathbf{n} = [n_1, \ldots, n_{m_2}]^{\mathrm{T}}$, then the vectors $\mathbf{f[k]} = [f[k_1], \ldots, f[k_{m_1}]]^{\mathrm{T}}$ and $\mathbf{g[n]} = [g[n_1], \ldots, g[n_{m_2}]]^{\mathrm{T}}$ contain the known signal values of f and g, respectively. Similarly, if the indices in vectors $\bar{\mathbf{k}}$ and $\bar{\mathbf{n}}$ represent unknown sample locations in $\mathbf{f}$ and $\mathbf{g}$, then $\mathbf{f}[\bar{\mathbf{k}}]$ and $\mathbf{g}[\bar{\mathbf{n}}]$ represent the unknown signal values.

Let $\mathbf{F^a[n,k]}$ be an $m_2 \times m_1$ submatrix of $\mathbf{F}^a$ obtained by choosing its n_1th,…,n_{m_2}th rows and k_1th,…,k_{m_1}th columns. By choosing the same rows and the remaining columns, one can also construct the submatrix $\mathbf{F^a[n,\bar{k}]}$, which is of size $m_2 \times (N - m_1)$. Then, the relation in Eq. (2) can be rewritten as

$$\mathbf{g[n]} = \mathbf{F^a[n,\bar{k}]f[\bar{k}]} + \mathbf{F^a[n,k]f[k]}. \tag{3}$$

Since only $\mathbf{f}[\bar{\mathbf{k}}]$ is unknown and required to be estimated in the above equation, the linear system of equations defining the forward problem is given by

$$\mathbf{g'} = \mathbf{F^a[n,\bar{k}]f[\bar{k}]}, \tag{4}$$

where $\mathbf{g'} = \mathbf{g[n]} - \mathbf{F^a[n,k]f[k]}$. Thus, in order to estimate $\mathbf{f}(\bar{\mathbf{k}})$, which contains the unknown signal values of $\mathbf{f}$, we need to solve

the above system of equations. Similarly, to infer $\mathbf{g}(\tilde{\mathbf{n}})$, which contains the unknown signal values of $\mathbf{g}$, one can solve

$$\mathbf{f}' = \mathbf{F}^{-\mathbf{a}}[\mathbf{k}, \tilde{\mathbf{n}}]\mathbf{g}(\tilde{\mathbf{n}}), \tag{5}$$

where $\mathbf{f}' = \mathbf{f}[\mathbf{k}] - \mathbf{F}^{-\mathbf{a}}[\mathbf{k}, \mathbf{n}]\mathbf{g}(\mathbf{n})$.

Now, notice that knowing the signal completely in one domain is equivalent to knowing it in all domains, since simple discrete fractional Fourier transformation allows us to compute the signal in any domain from that in any other. Physically this corresponds to the fact that knowing a wave field fully in any plane allows it to be computed anywhere else. Thus, when the signal is partially known in two domains, it is enough to estimate the signal either in the a_1th- or a_2th-order domain. We refer to the domain where we choose to estimate the signal as the *reference plane*. In this work, we choose the domain with the largest number of known samples as the reference plane. That is, if $m_1 \geq m_2$, the reference plane is chosen as the a_1th-order FRFD and we work with Eq. (4). Otherwise, it is chosen as the a_2th-order FRFD and we work with Eq. (5). We could also choose the reference plane as the domain with fewer number of samples or as any other domain that does not contain any known sample; however, our choice requires us to solve for a fewer number of unknowns and is therefore simpler.

Note that our problem formulation involves relating the known samples in two FRFDs to the unknown samples in these two domains. This formulation is general enough to be applicable to any two planes in a practical quadratic-phase optical system, which are systems involving an arbitrary sequence of lenses, sections of free space, and quadratic graded-index media. This is because quadratic-phase optical systems can be mathematically expressed in terms of linear canonical transforms (LCTs), and, based on the well-known Iwasawa decomposition [4], an arbitrary LCT can be decomposed into a FRT followed by scaling followed by chirp multiplication. Among these three operations (FRT, scaling, and chirp multiplication), the only operation that affects the conditioning of the inverse problem is the FRT operation. This is because (i) *scaling* is simply a reinterpretation of the samples (by scaling the sampling interval), and (ii) *chirp multiplication* can be performed through multiplication with a unitary matrix (a diagonal matrix whose diagonals are complex numbers with unit magnitude). Thus, these two operations do not affect the conditioning. As a result, the only important aspect of the two planes in terms of conditioning is the orders of the FRFDs associated with these planes, and these orders can be found by using the Iwasawa decomposition [4]. To summarize, a system involving an arbitrary sequence of lenses, sections of free space, and quadratic graded-index media, can be reduced to a FRT system, whose only parameter is the fractional order between the two domains $a = a_2 - a_1$. The condition numbers obtained will be the same for both systems. Thus the formulation we work with can be kept simple and tractable, while the results remain applicable to a broad class of systems of practical interest.

Before concluding this section, we will discuss a property of the discrete fractional Fourier transform that will be helpful in interpreting our numerical results. Again letting our signals be defined from $k = -N/2$ to $k = N/2 - 1$, consider an input vector $\mathbf{f}$ whose only nonzero element is at $k = 0$ with the value of unity. In other words, $\mathbf{f}$ is a discrete unit impulse centered at the origin. The discrete FRT of this vector is obtained by multiplying it by $\mathbf{F}^a$ and is the central column of the $\mathbf{F}^a$ matrix corresponding to $k = 0$. This vector is plotted for several values of a in Fig. 1.

Now, also recalling that larger values of a correspond to longer distances of propagation in free space, we see that the output corresponding to a centrally located impulse gets broader and broader as a increases from 0 to 1 and light propagates. Moreover, we can observe that the broadening is roughly proportionate to a. Further discussion may be found in Chap. 6 of [17]. If the input impulse is not centrally situated, the output will be more complicated due to edge and wraparound effects, but the same physical interpretation of broadening with increasing a remains.

What does this mean in terms of the conditioning of the system of equations? As the distance of propagation and the fractional order a increase, the response to an impulse broadens. For larger values of a, the broader response means that the input impulse *will have an effect on* a greater number of output points. Likewise, if we consider a given point at the output side, the larger the broadening, the value of light at that point *will be affected by* a greater number of input points. So larger a means that input or output points, either by affecting or being affected by, have interaction with a greater number of points on the other plane. This means they determine or depend on a greater number of points, which means they are less independent, or have more redundancy with respect to each other. In a setting containing such known samples, even if the matrix is not rank-deficient in a strict sense, some of its eigenvalues (or singular values) will be very small, making it effectively rank-deficient or ill-conditioned [33,40], leading to a large condition number and larger errors in the solution with small perturbations.

To be more specific, consider a known sample point on one of the two planes. (It does not matter which plane the point is on since both the FRT and the underlying physics of propagation are reciprocal.) Also consider, on the other plane, the extent of the response to an impulse at the location of the known point. This extent defines a region of sample points in the other domain that interact with the known sample point in the original domain. To a good approximation, the known

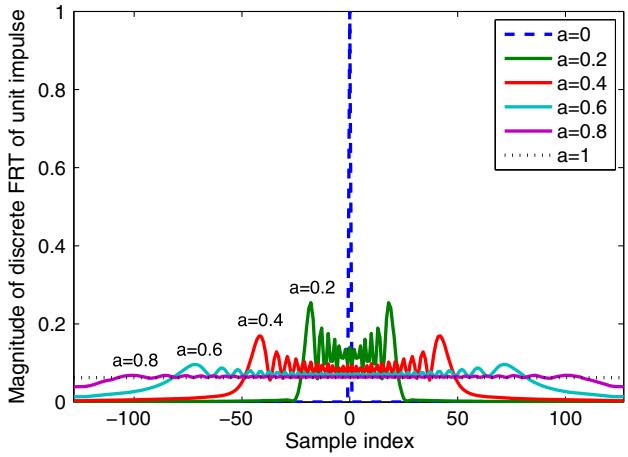

Fig. 1. Discrete FRT of a unit impulse centered at the origin for different values of a, for $N = 256$.

value can be expressed as a linear combination of the values at the points lying within its interaction region on the other plane. Now, if all the values within this region are already known, then the original known sample point will be totally redundant, since it could be calculated from the already known values. If, on the other hand, at least one value within this region is unknown, then the original known sample is not redundant. Thus, to avoid redundancy, whichever known value we choose, the extent on the other plane of the response to an impulse, should contain at least one unknown value. In fact, it is best if it contains precisely one. In this case, the known value in one plane directly determines the unknown in the other domain, essentially corresponding to a transfer of a known from one plane to another. However, when the interaction region of one known point overlaps with that of another known point, and the overlapping region contains one or more unknowns, things become more complicated. This means that the knowns on the original plane are dependent on common unknowns in the other plane.

In our numerical experiments, we investigate for different distributions of known information, the change in the condition number as a function of the order a. We consider the cases when the total number of known samples in the two domains is equal to or more than N. When interpreting the results, we will use the term *strict redundancy* to refer to the case when the system matrix has linearly dependent columns, or equivalently the matrix is rank-deficient. The term *effective redundancy* will refer to the case when the columns are close to being linearly dependent, but are not strictly so, or equivalently the matrix is effectively rank-deficient. The inverse problem is also effectively underdetermined in this case. For example, when we have linearly dependent rows, the associated known samples will be considered to be strictly redundant, whereas when the rows are close but not equal to being linearly dependent, the associated known samples will be considered to be effectively redundant. In our numerical experiments, we observed strict redundancy when $a = 0$ and $a = 1$. Since the discrete FRT matrix does not have a closed-form expression for other values of a, we can only speak of effective redundancy in these cases. We also note that the results may not be reliable when we are very close to the strict redundancy cases at $a = 0$ and $a = 1$, since chirp functions exhibit unusual behaviors in these limits.

3. ANALYSIS OF CONDITIONING

In this work, our aim is to analyze conditioning of the linear problem formulated in Eqs. (4) or (5). As well known, if there are less equations than unknowns, there cannot be a unique solution. For this reason, throughout this work, we focus only on the case where there are at least as many equations as unknowns; that is, the number of known samples in the two domains satisfies $m_1 + m_2 \geq N$. In this case, if the system is both full column rank and consistent, there exists a unique solution, but the accuracy of it depends on the conditioning of the problem. This is because the solution of the problem will be affected by limited machine precision, as well as inevitable model and sample errors, in a way that depends on its conditioning. The condition number of the system matrix determines how sensitive the solution is to such perturbations.

We will investigate how the condition number of $\mathbf{F}^{a}[\mathbf{n}, \bar{\mathbf{k}}]$ or $\mathbf{F}^{-a}[\mathbf{k}, \bar{\mathbf{n}}]$ is affected by the distribution and number of known samples. This will tell us how difficult it is for us to accurately recover the field from those known samples, since the condition number indicates how accurately the unknown samples can be estimated from the given samples.

The condition number associated with the linear equation $\mathbf{Ax} = \mathbf{b}$ gives a bound on the inaccuracy of the solution $\mathbf{x}$ and measures the rate at which the solution $\mathbf{x}$ will diverge from the true solution with respect to a change in $\mathbf{b}$ or $\mathbf{A}$. In other words, the condition number is an amplification factor that bounds the maximum relative error in the solution due to a given relative error in the input data when standard matrix inversion or least-squares techniques are used to obtain a solution [33,40]. The condition number of a nonsingular square matrix $\mathbf{A}$ is defined as [40]

$$\mathrm{cond}(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|, \tag{6}$$

with respect to a given matrix norm. For rectangular matrices, the pseudoinverse is used instead of the inverse in the definition of the condition number. The condition number of a matrix measures how close it is to being rank-deficient. For square matrices, it is a measure of how close it is to singularity.

In this study, the condition number in the spectral norm (induced by the vector 2-norm) is used, which is given by the ratio of the largest singular value of the matrix to the smallest [40]:

$$\mathrm{cond}_2(\mathbf{A}) = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}, \tag{7}$$

which is always ≥ 1, with equality satisfied when A is unitary. A condition number close to 1 indicates a well-conditioned matrix and small uncertainty in the solution. Note that the value of the condition number depends on the particular norm used; however, these values can differ by at most a fixed constant because of the equivalence of the underlying vector norms. Thus, our analysis is valid independent of the chosen matrix norm.

We will now present some important properties of the problem when $m_1 + m_2 = N$ with the system matrix being square and invertible. Using the fact that the matrix $\mathbf{F}^{a}$ is unitary and symmetric, we will show that all of the following matrices have the same 2-norm condition number: $\mathbf{F}^{a}[\mathbf{n}, \bar{\mathbf{k}}]$, $\mathbf{F}^{a}[\bar{\mathbf{k}}, \mathbf{n}]$, $\mathbf{F}^{-a}[\bar{\mathbf{k}}, \mathbf{n}]$, $\mathbf{F}^{-a}[\mathbf{n}, \bar{\mathbf{k}}]$. This indicates that the accuracy of the solution is the same for the following problems:

1. $\mathbf{F}^{a}[\mathbf{n}, \bar{\mathbf{k}}]$: The indices in $\mathbf{k}$ are known for the signal f and the indices in $\mathbf{n}$ are known for the signal g.
2. $\mathbf{F}^{a}[\bar{\mathbf{k}}, \mathbf{n}]$: The indices in $\bar{\mathbf{n}}$ are known for the signal f and the indices in $\bar{\mathbf{k}}$ are known for the signal g.
3. $\mathbf{F}^{-a}[\bar{\mathbf{k}}, \mathbf{n}]$: The indices in $\bar{\mathbf{k}}$ are known for the signal f and the indices in $\bar{\mathbf{n}}$ are known for the signal g.
4. $\mathbf{F}^{-a}[\mathbf{n}, \bar{\mathbf{k}}]$: The indices in $\mathbf{n}$ are known for the signal f and the indices in $\mathbf{k}$ are known for the signal g.

Thus, all of the above problems are equivalent to each other in terms of conditioning of the problem. If we consider the first and third ones together, it is clear that the accuracy of the solution is the same when we exchange the knowns and unknowns with each other in both domains. Moreover, if we consider the first and fourth ones together, changing the roles of the two

domains does not affect the accuracy of the solution. This supports the symmetrical structure of the problem and shows that, as expected from an optical interpretation, the direction of propagation does not create any difference on the solution.

The equivalence of the second problem in the list to the first one follows from the following derivation. The transpose of a matrix has the same singular values as the original matrix; therefore, we have $\mathrm{cond}_2(\mathbf{A}^{\mathrm{T}}) = \mathrm{cond}_2(\mathbf{A})$. Since $\mathbf{F}^a$ is a symmetric matrix, $\mathbf{F}^a[\mathbf{n}, \bar{\mathbf{k}}] = (\mathbf{F}^a[\bar{\mathbf{k}}, \mathbf{n}])^{\mathrm{T}}$. From the above result, we see that the condition numbers of $\mathbf{F}^a[\mathbf{n}, \bar{\mathbf{k}}]$ and $\mathbf{F}^a[\bar{\mathbf{k}}, \mathbf{n}]$ are the same. The equivalence of the third and fourth problems in the list follows similarly.

The equivalence of the fourth problem in the list to the first one follows from the fact that $\mathbf{F}^a$ is a unitary matrix: $(\mathbf{F}^a)^{\mathrm{H}} = \mathbf{F}^{-a}$ where $(\cdot)^{\mathrm{H}}$ denotes the Hermitian transpose of a matrix. Since $\mathbf{F}^a$ is also symmetric, $(\mathbf{F}^a)^* = \mathbf{F}^{-a}$, where $*$ denotes the conjugate of a matrix, and thus $(\mathbf{F}^a[\mathbf{n}, \bar{\mathbf{k}}])^* = \mathbf{F}^{-a}[\mathbf{n}, \bar{\mathbf{k}}]$. Since the conjugate of a matrix has the same singular values with the original matrix, we have $\mathrm{cond}_2(\mathbf{A}^*) = \mathrm{cond}_2(\mathbf{A})$. Then, $\mathbf{F}^a[\mathbf{n}, \bar{\mathbf{k}}]$ and $\mathbf{F}^{-a}[\mathbf{n}, \bar{\mathbf{k}}]$ have the same condition number. This completes the proof of the equivalence of all four problems in terms of conditioning.

As noted before, instead of choosing the domain with the largest number of known samples as the reference plane, one could also choose an empty domain as the reference plane. We will now discuss this case in more detail. Without loss of generality, we can say that the 0th-order FRFD does not contain any known sample, so that we can choose it as the reference plane. To formulate the problem, let us denote the signal in the reference plane as h. Then, the relation between the signals are given by $\mathbf{f} = \mathbf{F}^{\mathbf{a_1}}\mathbf{h}$ and $\mathbf{g} = \mathbf{F}^{\mathbf{a_2}}\mathbf{h}$. If we write these relations only for the known samples and combine everything together, we obtain

$$\begin{bmatrix} \mathbf{f}[\mathbf{k}] \\ \mathbf{g}[\mathbf{n}] \end{bmatrix} = \begin{bmatrix} \mathbf{F}^{a_1}[\mathbf{k}] \\ \mathbf{F}^{a_2}[\mathbf{n}] \end{bmatrix} \mathbf{h}, \tag{8}$$

where $\mathbf{F}^{a_1}[\mathbf{k}]$ and $\mathbf{F}^{a_2}[\mathbf{n}]$ denote submatrices of $\mathbf{F}^{a_1}$ and $\mathbf{F}^{a_2}$ obtained by choosing the k_1th,…,k_{m_1} th rows and the n_1th,…,n_{m_2} th rows, respectively. We need to solve the above system of equations in order to find the signal in the reference plane. Although $\mathbf{F}^{a_1}[\mathbf{k}]$ and $\mathbf{F}^{a_2}[\mathbf{n}]$ have orthogonal rows separately, their combination may contain non-orthogonal rows since the rows of FRT matrices of order a_1 and a_2 are not necessarily orthogonal to each other.

Based on the above formulation, we will now prove that the chosen reference plane does not create any difference on the conditioning of the problem. Let us choose the ath-order domain as reference. Then, the sensitivity of the solution depends on the condition number of the following matrix:

$$\mathbf{F_1} = \begin{bmatrix} \mathbf{F}^{a_1-a}(\mathbf{k}) \\ \mathbf{F}^{a_2-a}(\mathbf{n}) \end{bmatrix}. \tag{9}$$

If we instead choose the $\hat{a}$th-order domain as the reference such that $\hat{a} \neq a$, the sensitivity depends on the following matrix:

$$\mathbf{F_2} = \begin{bmatrix} \mathbf{F}^{a_1-\hat{a}}(\mathbf{k}) \\ \mathbf{F}^{a_2-\hat{a}}(\mathbf{n}) \end{bmatrix}. \tag{10}$$

Since $\mathbf{F}^{a_1-a}\mathbf{F}^{a-\hat{a}} = \mathbf{F}^{a_1-\hat{a}}$, if we multiply $\mathbf{F_1}$ from the right by $\mathbf{F}^{a-\hat{a}}$, we will obtain $\mathbf{F_2}$. That is, $\mathbf{F_2} = \mathbf{F_1}\mathbf{F}^{a-\hat{a}}$. Since $\mathbf{F}^{a-\hat{a}}$ is a unitary matrix and the spectral norm of a matrix is invariant under a unitary transformation, we have $\mathrm{cond}_2(\mathbf{F_2}) = \mathrm{cond}_2(\mathbf{F_1})$. Thus, the conditioning of the problem is the same for different choices of the reference plane.

This completes our discussion of the necessary mathematical issues so that we can now move on to presenting the results.

4. DESIGN OF NUMERICAL EXPERIMENTS

In the numerical results we present, we vary the fractional order a between the two domains over the range [0, 1] with a step size of 0.1 and the number of samples N in each domain is chosen as 256. We take m_1 and m_2, which represent the number of known samples in the two domains, as powers of 2 up to 128 and then choose their symmetric values with respect to 128, up to 256. For different m_1 and m_2 pairs, and different distributions of these known samples, the change in the logarithm of the condition number (to base 10) is investigated as a function of the order a. Remember that physically, an increase in a corresponds to an increase in the distance of propagation. Increasing the order a means a point in one domain will interact with (affect or be affected by) more points in the other domain.

We considered a number of potential scenarios illustrated in Fig. 2, where we denote known samples with dark squares and unknown samples with empty squares. There are two ways in which we categorize the distribution of the known samples: uniform versus accumulated distributions, and complementary versus overlapping distributions. In the uniform distribution, the known samples are distributed uniformly in both domains over the whole extent. In the accumulated distribution, the known samples are accumulated at one end in both domains. For both cases, the known sample locations in the two domains can either be complementing each other or overlapping with each other. For the complementary case, the samples that are unknown in one domain are known in the other domain, whereas for the overlapping case, the samples that are known in one domain are also known in the other domain. The four combinations are most easily understood with reference to Fig. 2, which has been drawn for the case where half of the

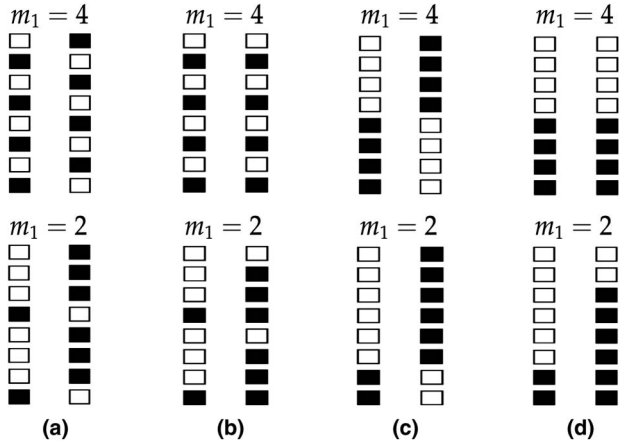

Fig. 2. Four basic known-sample distribution scenarios for $m_1 = 4$, $m_2 = 4$, and $m_1 = 2$, $m_2 = 6$ when $N = 8$. (a) Uniform-complementary, (b) uniform-overlapping, (c) accumulated-complementary, (d) accumulated-overlapping. The dark and empty squares denote the known and unknown samples in each domain, respectively.

samples are known in both domains ($m_1 = m_2$). When the number of known samples in the two domains are not equal to each other, the definition of complementary/overlapping distributions is such that, respectively, maximum possible non-overlapping/overlapping occurs between known samples in the two domains as illustrated in Fig. 2.

To give a more precise definition of these cases, if we denote the number of known samples in one domain as m, we use the following rules for the distributions of the known samples:

1. Uniform-overlapping distribution: For both domains:

   (a) For $m \leq N/2$, multiples of N/m are known and the remaining ones are unknown.

   (b) For $m > N/2$, the points which satisfy $N/(N-m) - 1 \pmod{N/(N-m)}$ are unknown and the remaining ones are known.

2. Uniform-complementary distribution: Rules for the a_1th-order domain are the same as above. For the a_2th-order domain, the rules are reversed as follows:

   (a) For $m \leq N/2$, the points which satisfy $N/m - 1 \pmod{N/m}$ are known and the remaining ones are unknown.

   (b) For $m > N/2$, multiples of $N/(N-m)$ are unknown and the remaining ones are known.

3. Accumulated-overlapping distribution: For both domains, the m samples with the most negative indices are known and the remaining ones are unknown.

4. Accumulated-complementary distribution: For the a_1th-order domain, the m samples with the most negative indices are known and the remaining ones are unknown whereas for the a_2th-order domain, the m samples with the most positive indices are known and the remaining ones are unknown.

The reason for considering these distributions is that they provide limiting cases that will allow us to draw general conclusions on how the distribution of known values affects our ability to recover the complete field. This in turn sheds light on which distributions contain known points that are more independent or more redundant with respect to each other. Moreover, these distributions may correspond to certain physical situations. For instance, the accumulated distributions correspond to situations where we know values only on one side of the field, or over a limited interval, and there is the need to complement this information with knowledge of the field at another plane. The uniform distribution corresponds to the physical situation where the known values are not spaced sufficiently closely together (with sufficient spatial resolution). In [41] the case of recovering the signal from two low-resolution versions at two different planes is considered, which corresponds to the uniform scenario here.

5. RESULTS

In what follows, we will consider the cases $m_1 + m_2 = N$ and $m_1 + m_2 > N$ separately.

A. Total Number of Knowns Equal to Number of Unknowns

First, we consider the case where the total number of known samples is equal to the number of unknowns: $m_1 + m_2 = N$. Figures 3 and 4 show the condition number versus a curves for accumulated and uniform distributions with the known samples shared differently between the two domains. For easier comparison of the different types of distributions, the curves for all different type of distributions are plotted together for the particular case of $m_1 = 8$, $m_2 = 248$, in Fig. 5.

As seen from the figures, as $m_1 \to N/2$, the condition number gets worse for all distributions. That is, the more equally the known samples are shared between the two domains, the more the condition number increases for all distributions. This indicates that distributing the N known samples equally to the two domains causes the largest amount of redundant information in the available data. To understand why, let us consider the domain with the greater number of known samples. Known samples in one domain are completely independent from each other, and there is no redundancy between them. Say we have 248 known samples in one domain and eight known samples in the other domain. Then these 248 known samples are strictly non-redundant since they are in the same domain. However, the other eight samples in the other domain will not be independent from the 248 samples since they are related to them through a propagation relationship. As $m_1 \to N/2$,

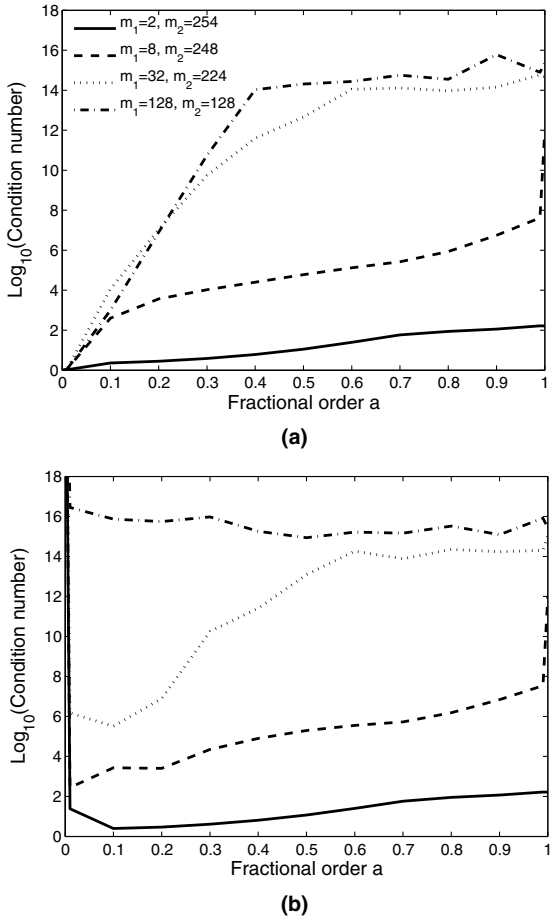

Fig. 3. Condition number versus a for complementary and overlapping accumulated distributions. (a) Accumulated-complementary distribution, (b) accumulated-overlapping distribution. The legend in part (a) indicating different pairs of m_1 and m_2, satisfying $m_1 + m_2 = N$, is valid for both plots.

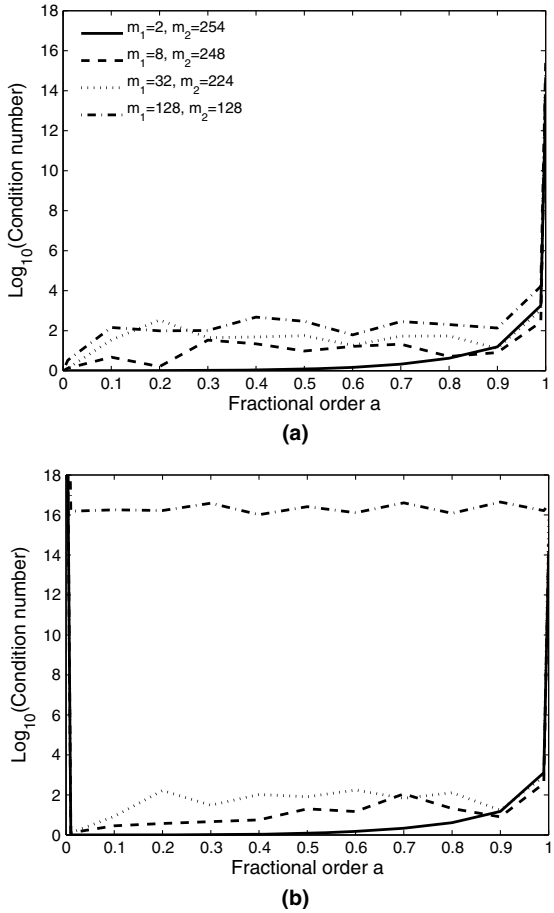

Fig. 4. Condition number versus a for complementary and overlapping uniform distributions. (a) Uniform-complementary distribution, (b) uniform-overlapping distribution. The legend in part (a) indicating different pairs of m_1 and m_2, satisfying $m_1 + m_2 = N$, is valid for both plots.

the number of strictly non-redundant known samples decreases. In fact, the larger of m_1 and m_2 is a lower bound for the number of non-redundant samples. As $m_1 \to N/2$, this lower bound goes to $N/2$.

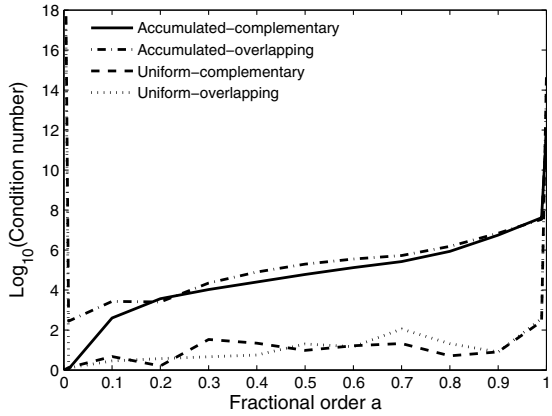

Fig. 5. Condition number versus a for all distributions when $m_1 = 8$ and $m_2 = 248$.

Observe from the figures that as the order a between two domains increases, the condition number also increases in general for all distributions. This is because, when a is small, the FRT of an impulse has a local broadening, causing only local dependency. On the other hand, as a approaches 1, the broadening covers the whole extent and each known point starts to depend on all points in the other domain. Hence, the region of interaction of each known sample broadens with increasing a. As a result, the region of interaction of different known samples also starts to overlap to a greater degree. This means that many known samples are at least partly related, and do not represent independent pieces of information. With this redundancy, knowledge of N values is actually worth less than knowledge of N *independent* values, and less sufficient to reliably determine the N unknown values.

Our observation that the condition number generally worsens with increasing order a has an important practical consequence. It is in general better to know samples from two planes whose fractional orders are close to each other (i.e., $a = a_2 - a_1$ is small). The physical location of such two planes, of course depends on the application and the optical system configuration. (The fractional orders corresponding to these physical locations can be found by virtue of the Iwasawa decomposition [4].) If free-space propagation is in question, this implies choosing planes that are physically close to each other. On the other hand, for an optical system involving lenses, two planes with fractional orders that are close to each other may as well be physically far from each other. In fact, multiple physical planes can serve the same purpose. This illustrates the generality of our formulation and its applicability to a rather broad class of optical systems.

As clearly seen from Fig. 5, uniform distributions give better condition numbers than the accumulated distributions, and the difference dramatically increases as a increases. Thus, known field values distributed more uniformly throughout a region carry more information than the same number of field values concentrated in a particular portion of that region. To see this, consider light emanating from known values in the domain with the fewer number of known samples. For very small values of a on the other domain, there is only one unknown inside the interaction region of a known point. That single unknown can be solved in terms of the known on the original plane; in other words, a known value can be transferred to the other plane and make up for missing knowns there. There is no dependency or redundancy in this case. However, as we increase a, also depending on the value of m_1, the known points on the original plane will start to have overlapping interaction regions that contain common unknowns. This will create effective redundancy between known samples. However, the increase in the condition number with increasing a is much less in the uniform case, compared to the accumulated case. This is because in the uniform case, the known samples are farther apart from each other, resulting in both less overlap between their interaction regions and fewer number of unknowns falling inside this overlap. Another way to understand the differing behavior between the uniform and accumulated cases is to note that the uniform case resembles an interpolation problem whereas the accumulated case resembles an extrapolation problem, the latter of

which is generally recognized to be more difficult to be solved numerically. In the uniform case, there is strict redundancy at $a = 1$ since the matrix is explicitly rank-deficient in this case (see the appendix in Chap. 7 of [17] for a proof). Thus, the numerically obtained values displayed in the figures for $a = 1$ are not very meaningful and represent numerical limitations; theoretically they are infinite for this case.

Having discussed the general dependence on a, the effect of increasing m_1, and the different behavior of the uniform and accumulated cases, let us now compare the complementary and overlapping cases. We observe that while the complementary case is superior to the overlapping case for smaller values of a, for larger values of a they give similar results. The reason for this is that when the domains are close to each other, each known point is interacting with a small number of samples in the other domain since the extent of the response region for each point is small. Therefore, as a consequence of complementarity, most of the points lying in these regions of interaction are unknown samples, so that there is less redundancy. Indeed, it is instructive to compare the complementary and overlapping cases when $a = 0$. The complementary case reduces to the case of knowing each and every sample in a given plane, allowing perfect recovery. The overlapping case, on the other hand, corresponds to the case of twofold redundancy of all known samples along with zero knowledge of the remaining samples, a catastrophic case of redundancy. Since it is reasonable to expect this behavior to be inherited to some extent by small nonzero values of a, the observed behavior is not surprising. However, as the value of a is further increased, the regions of interaction become larger, and a comparable number of known sample points fall into these regions in both the complementary and overlapping cases, causing a comparable degree of redundancy in both cases. Thus we conclude that for values of a that are not small, shifting the known sample values in the transverse direction has little effect on their collective information value.

The accumulated-overlapping case deserves additional scrutiny. In this case, for values of m_1 that are close to $N/2$ (more precisely, for $m_1 > 32$), the condition number hits the MATLAB precision limit ceiling around 10^{16}, for all values of the fractional order a. We obtain condition numbers that do not hit the precision limit ceiling only for considerably uneven distributions (more precisely, for $m_1 \leq 32$). As we have discussed before, the redundancy is related to the maximum of

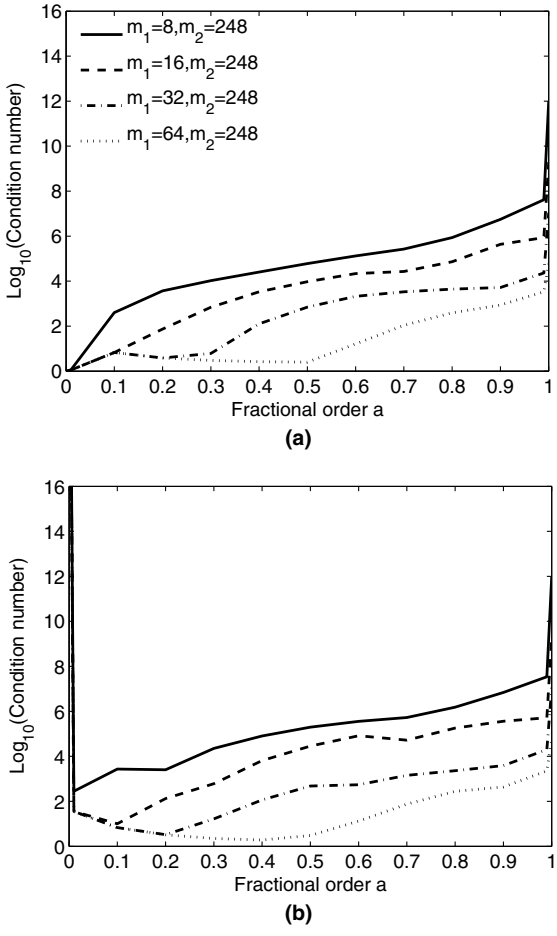

Fig. 6. Condition number versus a for complementary and overlapping accumulated distributions. (a) Accumulated-complementary distribution, (b) accumulated-overlapping distribution. The legend in part (a) indicating different pairs of doubling m_1 and constant m_2, satisfying $m_1 + m_2 \geq N$, is valid for both plots.

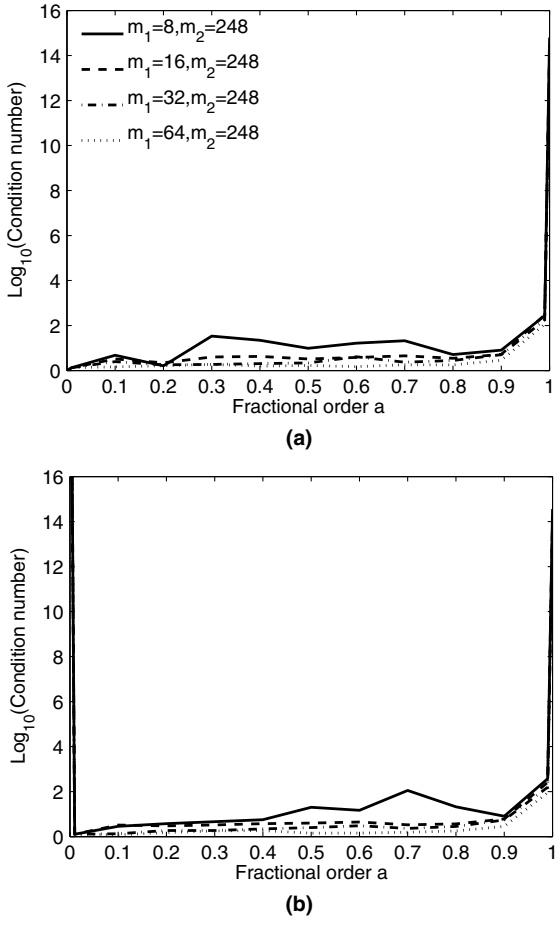

Fig. 7. Condition number versus a for complementary and overlapping uniform distributions. (a) Uniform-complementary distribution, (b) uniform-overlapping distribution. The legend in part (a) indicating different pairs of doubling m_1 and constant m_2, satisfying $m_1 + m_2 \geq N$, is valid for both plots.

m_1 and m_2. For the $a = 0$ case, we have strict redundancy. If the distribution is considerably uneven (m_1 is very different than m_2), then the maximum of m_1 and m_2 will still be close to N. For this reason, increasing a even slightly above 0, the system departs from strict redundancy quickly, since the additional information provided by the known samples in the other domain can compensate for the information gap. However, for distributions where m_1 is close to m_2, the maximum of m_1 and m_2 is much smaller than N, and thus the additional information from other known samples cannot close this information gap since the relative dependency between these points is high, as discussed before. It is also interesting to note that there often exists a FRT order between 0 and 1 that makes the condition number minimum, and this optimal FRT order depends on the balance between m_1 and m_2.

Finally, let us take another look at the accumulated-overlapping distribution. Note that since the discrete FRT has a periodic nature, the effect of a sample at the edge of the field wraps around to the other edge, which leads the system to exhibit complementary-like behavior to some degree. Of course, this behavior, which is a consequence of discretization, does not exist in the continuous setting. Therefore, the actual results corresponding to the real physical situation may be expected to be even worse than portrayed here.

B. Total Number of Knowns More Than Number of Unknowns

Here we investigate the improvement in the condition number when $m_1 + m_2 > N$, or equivalently, when we increase the total number of known samples beyond N. The curves in Figs. 6 and 7 are obtained by starting with the $m_1 = 8$ and $m_2 = 248$ case and doubling m_1 each time.

We observe that as we increase the number of known samples in one domain, the condition numbers improve as expected. The improvement is more dramatic in the accumulated case since the condition numbers there were much larger to begin with, and less pronounced in the uniform case since the condition numbers there were already not very large. Also observe that for accumulated distributions, even when the total number of samples exceeds N considerably, the condition number remains high for large values of a. The behavior of the curves can be understood through similar arguments as in the $m_1 + m_2 = N$ case. For instance, for the accumulated-complementary case, the sharp increase in the condition number starts when the extent of the response region associated with

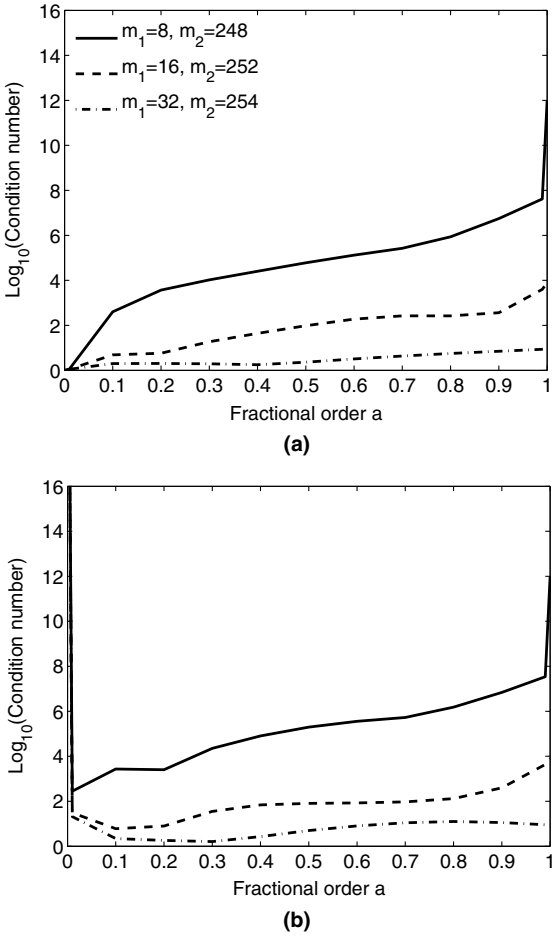

Fig. 8. Condition number versus a for complementary and overlapping accumulated distributions. (a) Accumulated-complementary distribution, (b) accumulated-overlapping distribution. The legend in part (a) indicating different pairs of doubling m_1 and increasing m_2, satisfying $m_1 + m_2 \geq N$, is valid for both plots.

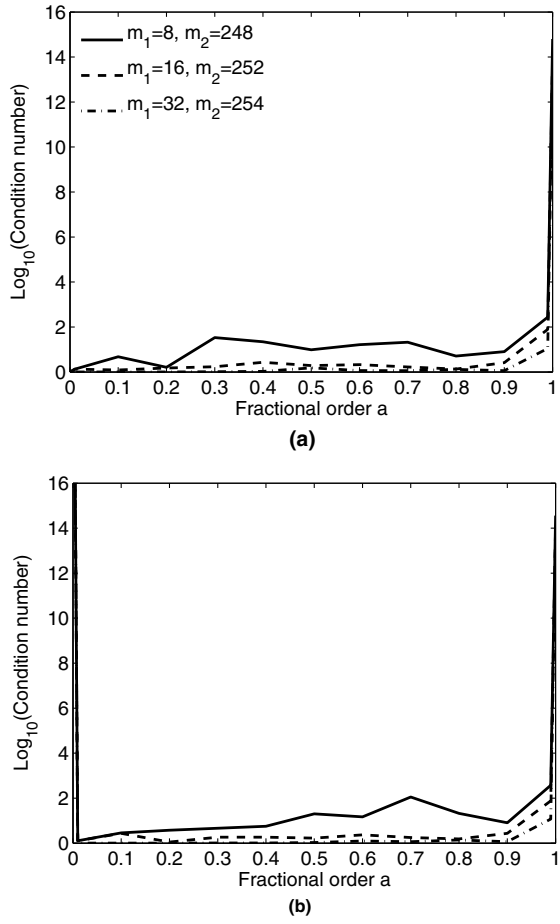

Fig. 9. Condition number versus a for complementary and overlapping uniform distributions. (a) Uniform-complementary distribution, (b) uniform-overlapping distribution. The legend in part (a) indicating different pairs of doubling m_1 and increasing m_2, satisfying $m_1 + m_2 \geq N$, is valid for both plots.

an unknown sample contains at least one unknown in the other domain, since in this case the known values inside this region no longer directly determine the unknown in the original domain.

As an alternative to keeping m_2 constant, we modified the way m_1 and m_2 change so that m_2 also increases while m_1 doubles. As seen from Figs. 8 and 9, this improves the condition number considerably for the accumulated case. There is improvement in the uniform case as well, but less pronounced since the condition numbers are much smaller to begin with. For the accumulated case, the increase in the condition number with increasing a is observed at a larger value of a, since the extent of the response region of an unknown sample starts to contain an unknown at larger values of a. Moreover, the minimum condition number over a is lower for all m_1 and m_2 pairs. This is because the greater of m_1 and m_2 gives us a basic minimum number of samples whose independence is guaranteed, and this increases when we increase m_1 and m_2.

Last, we investigate the case when there are an equal number of samples in both domains ($m_1 = m_2$). As seen from Figs. 10 and 11, the condition number is very small for most cases where $m_1 + m_2 > N$. The accumulated-overlapping case is a striking exception, where even considerably large total number

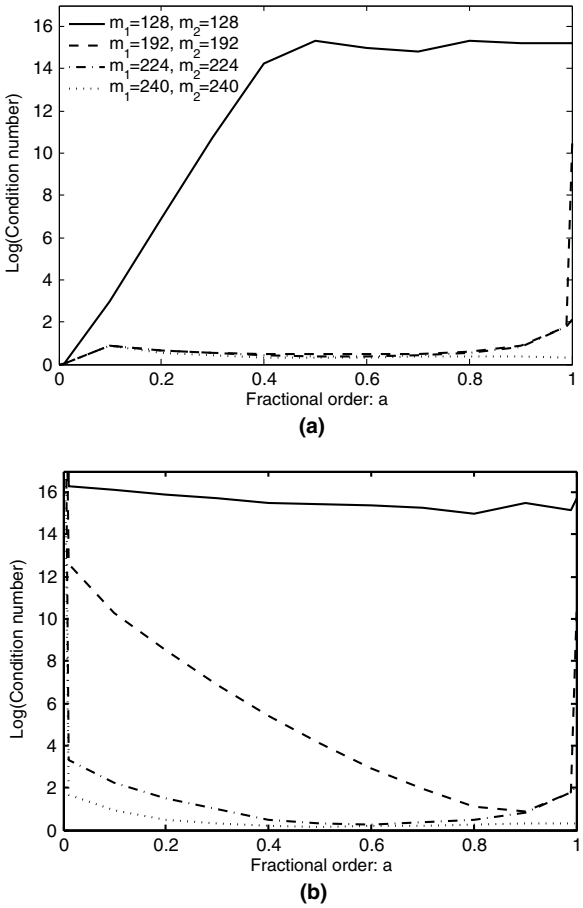

Fig. 10. Condition number versus a for complementary and overlapping accumulated distribution. (a) Accumulated-complementary distribution, (b) accumulated-overlapping distribution. The legend in part (a) indicating different pairs of $m_1 = m_2$, satisfying $m_1 + m_2 \geq N$, is valid for both plots.

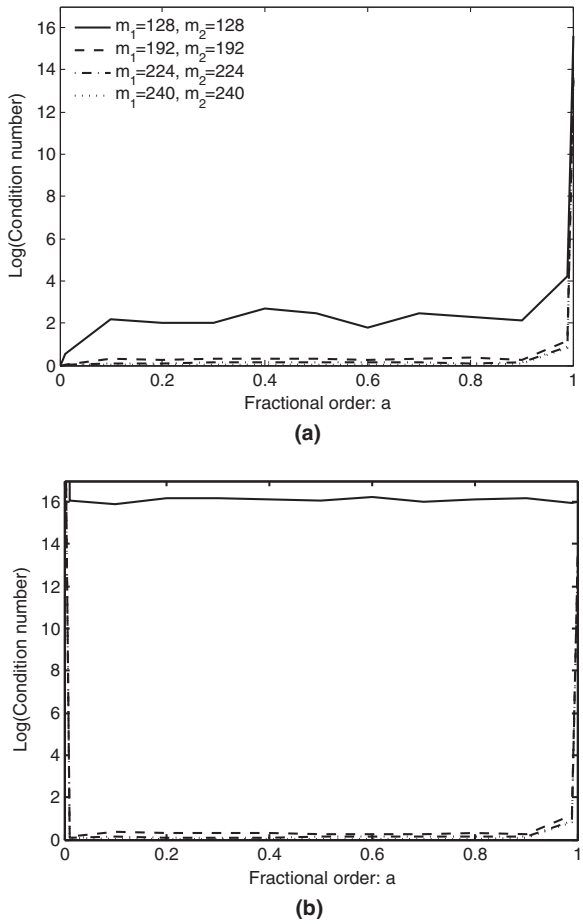

Fig. 11. Condition number versus a for complementary and overlapping uniform distributions. (a) Uniform-complementary distribution, (b) uniform-overlapping distribution. The legend in part (a) indicating different pairs of $m_1 = m_2$, satisfying $m_1 + m_2 \geq N$, is valid for both plots.

of known samples do not give small condition numbers for smaller values of a. That is, the system of equations cannot quickly depart from the strict redundancy case at $a = 0$. This different behavior with respect to a, compared to the $m_1 + m_2 = N$ cases, deserves explanation. For small values of a, there are unknowns which do not fall within the response region of any of the knowns in the other domain. Thus, roughly speaking, all of the knowns in one domain say nothing about some of the unknowns in the other domain, hindering our ability to solve for them. This situation is overcome when every unknown is covered by the extent of a response region originating from at least one known. The condition number improves quickly up to this point, with only small changes occurring thereafter.

6. CONCLUSION

We considered the mathematical problem of recovering a signal from partial and redundant information distributed over two fractional Fourier transform domains. Physically, this corresponds to the problem of recovering an optical field from partial and redundant information distributed over two planes

perpendicular to the optical axis. We took a novel linear algebraic approach to this problem and used the condition number as a measure of redundant information in the available sample values. By analyzing the effect of the number of known samples and their distributions on the condition number, we have explored the redundancy and information relations between the given data under different partial information conditions. This led us to a number of general observations on how the distribution of samples affects our ability to accurately solve for the optical field. As was very evident from our numerical experiments, the distribution of known samples over space was at least as important as their number in determining the solution.

We now summarize our main observations. But first recall that when we have an impulse in one plane, its discrete fractional Fourier transform on the other plane (corresponding to the light that would propagate from that point) has an extent that increases with increasing order a (corresponding to increasing distance of propagation). This extent defines a region of sample points in the other domain that interacts with a sample point at the original domain where the impulse is. With the term "interact," we mean that those sample points are related to each other through a linear expression involving coefficients that are substantially different than zero. Thus the larger the order a between the two planes, the greater the number of points on the other plane that will interact with a given point in the original plane. For small values of a, there is only local dependency but as a goes to unity, there is dependence on the whole transverse field.

We saw that when the known samples are shared more equally between the two domains, we have greater redundancy between these known samples, the condition number is worse, and it is much more difficult to recover the field accurately. It is more advantageous for more of the known samples to be in one plane and best when all of them are in the same plane, in which perfect determination of the field everywhere is p ossible. The critical quantity here is the greater of the number of knowns in the two planes. The larger this quantity is the better, since it gives us a guaranteed number of samples that are independent. It is worst when the number of knowns are shared equally between the two planes. The practical consequence of this observation in the design of an optical system is that one should try to choose as many samples as possible from one plane, and then compensate for the missing ones by taking additional samples from a second plane.

Generally speaking, as the fractional order between the domains increases, the condition number and the quality of the solution gets worse. This physically corresponds to increasing the distance between the transverse planes in free space, which leads to greater interaction, dependence, and therefore redundancy between the known samples, as a result of the broadening of the response to an impulse with increasing distance. The practical design consequence of this observation is that one should choose the two planes to correspond to FRFDs of close orders. Moreover, as occurred with the accumulated-overlapping distribution, there is sometimes an optimal fractional order, in which case the two planes should be chosen so that the difference of their fractional orders is this optimal value.

Having the known sample values in both domains uniformly distributed throughout the transverse extent of the field is better than having them huddled up in one side or region since in the latter case, the proximity of the known samples leads to greater overlap of their regions of interaction in the other domain, leading to greater dependency and redundancy. Another way of understanding this is to note that a more uniform distribution makes the problem more akin to interpolation, which is usually better posed than the extrapolation problem which arises when the samples are accumulated to a region. This suggests that we should choose equally spaced samples distributed over the support of the signal, rather than taking finer samples from a limited region.

Unless the fractional order between the two domains is very small, it does not make much difference whether the known samples are on directly corresponding grid points (overlapping) or alternating grid points (complementary). This is because unless the two planes are very close, shifting the sample points on the other side has little effect on what falls into the relatively broad regions of interaction. However, for small values of a (corresponding to small physical distance between the two planes in free space), having the known sample locations in the two domains complement each other rather than overlap with each other is clearly better, as evident by considering the limiting case of $a = 0$ where the complementary case corresponds to perfectly independent known sample values while the overlapping case corresponds to a case of extreme redundancy.

Finally, increasing the total number of known samples beyond the number of unknowns is generally helpful. This not only may increase the greater of m_1 and m_2, and hence the guaranteed number of samples that are independent, but even when that does not change, it can reduce the condition number considerably. In this overdetermined case, extra information beyond the number of unknowns we are trying to solve for compensates for the information redundancy between the known samples. The benefit of increasing the total number of samples is less, however, when the condition numbers are already satisfactory.

Note that all these observations regarding the effects of the parameters and various different cases are consistent with the fact that the ideal scenario is to have all known samples in the same domain (on the same plane), in which case they are all independent, leading to perfect recovery. All the other cases may be considered as departures from this ideal situation.

In conclusion, we have observed how the spatial distribution of a number of known samples affects our ability to recover an optical field accurately. Certain qualities of this distribution, as we have summarized in the above paragraphs, reduce the value of given samples in recovering the field by making them more dependent and hence redundant. This dependency, of course, stems from the basic laws of propagation governing how the sample points at different planes are related to each other.

Acknowledgment. H. M. Ozaktas was supported in part by Türkiye Bilimler Akademisi (TÜBA).

REFERENCES

1. D. Mendlovic and H. M. Ozaktas, "Fractional Fourier transforms and their optical implementation: I," J. Opt. Soc. Am. A **10**, 1875–1881 (1993).
2. A. W. Lohmann, "Image rotation, Wigner rotation, and the fractional Fourier transform," J. Opt. Soc. Am. A **10**, 2181–2186 (1993).
3. H. M. Ozaktas, M. A. Kutay, and C. Candan, "Fractional Fourier transform," in *Transforms and Applications Handbook*, A. D. Poularikas, ed. (CRC Press, 2010), pp. 14-1–14-28.
4. F. S. Oktem and H. M. Ozaktas, "Equivalence of linear canonical transform domains to fractional Fourier domains and the bicanonical width product: a generalization of the space-bandwidth product," J. Opt. Soc. Am. A **27**, 1885–1895 (2010).
5. H. Ozaktas, S. Arık, and T. Coşkun, "Fundamental structure of Fresnel diffraction: natural sampling grid and the fractional Fourier transform," Opt. Lett. **36**, 2524–2526 (2011).
6. H. Ozaktas, S. Arık, and T. Coşkun, "Fundamental structure of Fresnel diffraction: longitudinal uniformity with respect to fractional Fourier order," Opt. Lett. **37**, 103–105 (2012).
7. K. B. Wolf, "Construction and properties of canonical transforms," in *Integral Transforms in Science and Engineering* (Plenum, 1979), Chap. 9.
8. H. M. Ozaktas and M. F. Erden, "Relationships among ray optical, Gaussian beam, and fractional Fourier transform descriptions of first-order optical systems," Opt. Commun. **143**, 75–86 (1997).
9. D. P. Kelly, B. M. Hennelly, W. T. Rhodes, and J. T. Sheridan, "Analytical and numerical analysis of linear optical systems," Opt. Eng. **45**, 088201 (2006).
10. H. M. Ozaktas, A. Koç, I. Sari, and M. A. Kutay, "Efficient computation of quadratic-phase integrals in optics," Opt. Lett. **31**, 35–37 (2006).
11. M. J. Bastiaans and T. Alieva, "Classification of lossless first-order optical systems and the linear canonical transformation," J. Opt. Soc. Am. A **24**, 1053–1062 (2007).
12. A. Koc, H. M. Ozaktas, C. Candan, and M. A. Kutay, "Digital computation of linear canonical transforms," IEEE Trans. Signal Process. **56**, 2383–2394 (2008).
13. F. S. Oktem and H. M. Ozaktas, "Exact relation between continuous and discrete linear canonical transforms," IEEE Signal Process. Lett. **16**, 727–730 (2009).
14. H. M. Ozaktas and F. S. Oktem, "Phase-space window and degrees of freedom of optical systems with multiple apertures," J. Opt. Soc. Am. A **30**, 682–690 (2013).
15. F. S. Oktem and H. M. Ozaktas, "Linear canonical domains and degrees of freedom of signals and systems," in *Linear Canonical Transforms* (Springer, 2016), pp. 197–239.
16. S. Ö. Arık and H. M. Ozaktas, "Optimal representation and processing of optical signals in quadratic-phase systems," Opt. Commun. **366**, 17–21 (2016).
17. F. S. Oktem, "Signal representation and recovery under partial information, redundancy, and generalized finite extent constraints," Master's thesis (Bilkent University, 2009).
18. F. S. Oktem and H. M. Ozaktas, "Linear algebraic analysis of fractional Fourier domain interpolation," in *IEEE 17th Signal Processing and Communications Applications Conference*, Antalya, Turkey, 2009, pp. 873–875.
19. F. S. Oktem and H. M. Ozaktas, "Condition number in recovery of signals from partial fractional Fourier domain information," in *Applied Industrial Optics: Spectroscopy, Imaging and Metrology* (Optical Society of America, 2013).
20. N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast Fourier transform approach," J. Opt. Soc. Am. A **15**, 857–867 (1998).
21. K. Matsushima, H. Schimmel, and F. Wyrowski, "Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves," J. Opt. Soc. Am. A **20**, 1755–1762 (2003).
22. L. Onural and H. M. Ozaktas, "Signal processing issues in diffraction and holographic 3DTV," Signal Process. **22**, 169–177 (2007).
23. L. Onural, A. Gotchev, H. M. Ozaktas, and E. Stoykova, "A survey of signal processing problems and tools in holographic three-dimensional television," IEEE Trans. Circuits Syst. Video Technol. **17**, 1631–1646 (2007).
24. G. B. Esmer, V. Uzunov, L. Onural, H. M. Ozaktas, and A. Gotchev, "Diffraction field computation from arbitrarily distributed data points in space," Signal Process. **22**, 178–187 (2007).
25. H. M. Ozaktas and L. Onural, *Three-Dimensional Television: Capture, Transmission, Display*, Springer Series in Signals and Communication Technology (Springer, 2008).
26. G. B. Esmer, L. Onural, and H. M. Ozaktas, "Exact diffraction calculation from fields specified over arbitrary curved surfaces," Opt. Commun. **284**, 5537–5548 (2011).
27. H. E. Guven, H. M. Ozaktas, A. E. Cetin, and B. Barshan, "Signal recovery from partial fractional Fourier domain information and its applications," IET Signal Process. **2**, 15–25 (2008).
28. A. Özçelikkale, H. M. Ozaktas, and E. Arikan, "Signal recovery with cost-constrained measurements," IEEE Trans. Signal Process. **58**, 3607–3617 (2010).
29. A. Özçelikkale and H. M. Ozaktas, "Representation of optical fields using finite numbers of bits," Opt. Lett. **37**, 2193–2195 (2012).
30. A. Özçelikkale and H. M. Ozaktas, "Beyond Nyquist sampling: a cost-based approach," J. Opt. Soc. Am. A **30**, 645–655 (2013).
31. T. C. Gulcu, "Finite representation of finite energy signals," Master's thesis (Bilkent University, 2011).
32. P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms* (SIAM, 2010), Vol. **7**.
33. P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion* (SIAM, 1998), Vol. **4**.
34. C. R. Vogel, *Computational Methods for Inverse Problems* (SIAM, 2002), Vol. **23**.
35. M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging* (CRC Press, 1998).
36. A. Ribes and F. Schmitt, "Linear inverse problems in imaging," IEEE Signal Process. Mag. **25**(4), 84–99 (2008).
37. X.-G. Xia, "On bandlimited signals with fractional Fourier transform," IEEE Signal Process. Lett. **3**, 72–74 (1996).
38. A. Zayed, "On the relationship between the Fourier and fractional Fourier transforms," IEEE Signal Process. Lett. **3**, 310–311 (1996).
39. C. Candan, M. Kutay, and H. Ozaktas, "The discrete fractional Fourier transform," IEEE Trans. Signal Process. **48**, 1329–1337 (2000).
40. M. T. Heath, *Scientific Computing: An Introductory Survey* (McGraw-Hill, 2002).
41. A. E. Cetin, H. Özaktaş, and H. M. Ozaktas, "Resolution enhancement of low resolution wavefields with POCS algorithm," Electron. Lett. **39**, 1808–1810 (2003).