Optimized Least-Square Nonuniform Fast Fourier Transform

Mathews Jacob, *Member, IEEE*

***Abstract*—The main focus of this paper is to derive a memory efficient approximation to the nonuniform Fourier transform of a support limited sequence. We show that the standard nonuniform fast Fourier transform (NUFFT) scheme is a shift invariant approximation of the exact Fourier transform. Based on the theory of shift-invariant representations, we derive an exact expression for the worst-case mean square approximation error. Using this metric, we evaluate the optimal scale-factors and the interpolator that provides the least approximation error. We also derive the upper-bound for the error component due to the lookup table based evaluation of the interpolator; we use this metric to ensure that this component is not the dominant one. Theoretical and experimental comparisons with standard NUFFT schemes clearly demonstrate the significant improvement in accuracy over conventional schemes, especially when the size of the uniform fast Fourier transform (FFT) is small. Since the memory requirement of the algorithm is dependent on the size of the uniform FFT, the proposed developments can lead to iterative signal reconstruction algorithms with significantly lower memory demands.**

***Index Terms*—Fourier transform, interpolation, nonuniform, sampling, shift-invariant.**

I. Introduction

THE efficient evaluation of the nonuniform Fourier samples of an N-point discrete signal is a central problem in many areas including tomography [1], [2], magnetic resonance imaging [3]–[5], synthetic aperture radar [6], recovery of band-limited images [7], [8], filter design [9], [10], and wavelets [11]. The brute-force evaluation of the Fourier sum at nonuniform signal samples is computationally expensive. Hence, the standard approach is to obtain these samples as the interpolation of the uniform Fourier transform. The uniform Fourier transform is computed using the standard K-point fast Fourier transform (FFT) ($K \geq N$), while support limited functions (e.g., Kaiser–Bessel, Gaussian) are used to interpolate the samples. It is reported that weighting the signal with suitable scale-factors, before evaluating the uniform FFT, significantly reduces the approximation error [12], [13]. This approach is often referred to as "type-2" nonuniform fast Fourier transform (NUFFT), to differentiate it from the gridding scheme (type-1 NUFFT) [12]–[15].

Our work is motivated by iterative non-Cartesian MRI [3]–[5] and iterative tomographic reconstruction [1], [2]. NUFFT schemes are often used to accelerate the computation of the forward model, which needs to be evaluated during each iteration [4], [5]. Current NUFFT schemes have some drawbacks that limit their practical utility in reconstructing large multidimensional datasets. Since the accuracy of the reconstruction algorithms are heavily dependent on the quality of the NUFFT approximation, it is a general practice to evaluate the Fourier transform on a fine uniform grid (e.g., $K = 2N$) to minimize the interpolation error [13]. This approach significantly increases the memory demands of the algorithm. For example, the reconstruction of a three dimensional data-set with $K = 2N$ requires eight times more memory that the original data-set. Another drawback is in the suboptimal selection of scale factors. Although scale-factors play a significant role in reducing the NUFFT error, they are often selected arbitrarily [12] or are restricted to parametric families with few degrees of freedom [13]. This limits the performance of the NUFFT approximation significantly. Finally, there is no well-accepted methodology for selecting the sampling step in the discretization of the interpolators themselves. The exact evaluation of the interpolators at noninteger samples, within the iteration loop, is computationally expensive. Hence, it is a general practice to obtain these samples by linearly interpolating their precomputed uniform samples. This practice may considerably increase the approximation error, if the sampling step is not properly selected.

The main focus of this paper is to overcome the above limitations and thus derive a memory efficient approximation to the nonuniform Fourier transform. We show that the widely used NUFFT scheme is essentially a periodic shift invariant approximation [16]–[18] of the exact discrete Fourier transform. Based on our earlier results [19], we derive an exact expression for the worst-case mean square approximation error. This metric conveniently decouples the error contributions resulting from the scale-factors and the interpolator into two separate positive terms. This enables us to optimize both the scale-factors and the interpolator using the same performance measure. Specifically, we obtain a closed form expression for the optimal least-square scale-factors (OLS scale-factors) for a specified interpolator. Assuming these scale-factors, we derive the error metric that is only dependent on the interpolator. We then introduce an iterative re-weighted minimization algorithm to obtain the optimized least-square interpolator (OLS-interpolator). Using analogous arguments, we also derive the worst-case error resulting from

Manuscript received April 24, 2008; accepted December 25, 2008. First published February 06, 2009; current version published May 15, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Pierre Vandergheynst. This work is supported by the Clinical and Translational Research Institute at the University of Rochester.

M. Jacob is with the Department of Biomedical Engineering, University of Rochester, NY 14622 USA (e-mail: mathews.jacob@rochester.edu; website: http://www.cbig.rochester.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2014809

the discretization (lookup table based evaluation) of the interpolator. This expression is useful in determining the minimum oversampling factor that is required so that the performance of the interpolator is not limited by discretization. From theoretical and numerical comparisons, we find that the OLS-NUFFT significantly ameliorates the accuracy over classical approximations, especially when the length of the uniform FFT is small. Since the length of the uniform FFT determines the memory demands of the algorithm, these developments can lead to a memory efficient multidimensional NUFFT scheme.

Our paper is inspired by the work of Fessler *et al.* where the authors derived the interpolator by optimizing the maximum worst-case NUFFT approximation error [13]. An analogous expression was derived by Neislony *et al.* using the matrix formulation [20]. The main limitation of these expressions is their inability to decouple the effect of the scale-factors from that of the interpolation function. Hence, Fessler *et al.* restricted the scale-factors to a parametric family with few degrees of freedom and derived optimal parameters using exhaustive search [13]. An alternative is to assume Kaiser–Bessel or Gaussian scale-factors and optimize the interpolator [13], [20]. Since these approaches restricted the flexibility of the scale-factors and hence the optimal interpolation functions, the performance of these NUFFT schemes was only comparable to the one using the min–max optimized Kaiser-Bessel function [13]. Both of the above schemes ignored the error in discretizing the interpolation function. We find that this is a dominant component, often limiting the performance of the NUFFT scheme.

Error metrics, analogous to the proposed one, were proposed by several authors in the context of NUFFT. Steidl *et al.* developed an expression for the L_∞ norm of the approximation error [15], [21]. This expression also decouples the effect of the interpolator and scale-factors into two terms. However, it is not obvious on how to use this expression to derive the optimal interpolator and scale-factors since it is significantly more nonlinear than the proposed one. In the context of gridding (type 1 NUFFT), Jackson *et al.* used the out of band energy of the interpolation function as a measure of the aliasing error [22], while Beatty *et al.* developed a related expression that quantified the point wise aliasing amplitude [23]. Although these expressions were derived in the context of gridding, they bear some similarity to the first term of our band-limited metric (for almost band-limited functions). Jackson *et al.* used an iterative scheme to derive a finitely supported interpolating function that is maximally band-limited [22]; the resulting function is an approximation for the prolate spheroidal wave function. In contrast, Beatty *et al.* used second order cone programming to derive the interpolator that minimized the point-wise aliasing amplitude [23]. They obtained a slight improvement in aliasing error over the Kaiser–Bessel function, which in-turn is an approximation of the prolate spheroidal function.

The rest of the paper is organized as follows. In the next section, we review the fundamentals of periodic shift invariant representation, its error analysis, and the standard NUFFT method. In Section III, we apply the error expression to the NUFFT method and derive the optimal least-square scale-factors and outline the derivation of the interpolator. In the following section, we derive the optimal discretization of the interpolator and quantify the error contribution resulting from the lookup table based evaluation of the interpolator. We compare the performance of the interpolator to the standard methods using theoretical metrics as well as numerical simulations in Section V.

II. Preliminaries

A. Shift Invariant Approximation of Periodic Signals

The shift invariant representation is widely used for functions in $L_2(\mathbb{R})$ [16]–[18]. The shift invariant representation of the signal $s(\nu)$ is given by

$$s(\nu) = \sum_{k=-\infty}^{\infty} c[k]\varphi\left(\frac{\nu}{T} - k\right) \tag{1}$$

where $\varphi \in L_2(\mathbb{R})$ is the shift invariant basis function and $c[k] \in \ell_2(\mathbb{Z})$ are the coefficients. The basis function is often assumed to be finitely supported in $[-J/2, J/2]$, where J is an integer. This scheme can be extended to N-periodic signals, when the period N is an integer multiple of the sampling step (i.e., $N = KT$) [19], [24]. We assume N and K to be even integers, although there is no such restriction in the general setting [19]. For periodic signals $(s(\nu) = s(\nu + lN); \forall l \in \mathbb{Z})$, we have

$$\sum_{k=-\infty}^{\infty} c[k]\varphi\left(\frac{K}{N}\nu - k\right) = \sum_{k=-\infty}^{\infty} c[k]\varphi\left(\frac{K}{N}\nu - k + lK\right);\ \forall l \in \mathbb{Z}. \tag{2}$$

Assuming $\{\varphi(\nu/T - k), k \in \mathbb{Z}\}$ to be a linearly independent set of basis functions, this relation is only satisfied when $c[k + lK] = c[k], \forall l$. Using this K-periodicity of $c[k]$, we rewrite the above equation as

$$s(\nu) = \sum_{k=-K/2}^{K/2-1} c[k]\varphi_p\left(\frac{K}{N}\nu - k\right) \tag{3}$$

where φ_p is the K-periodized version of $\varphi(\nu) \in L_2(\mathbb{R})$:

$$\varphi_p(\nu) = \sum_{k\in\mathbb{Z}} \varphi(\nu - kK). \tag{4}$$

The coefficients $c[k]; k = K/2 \ldots K/2 - 1$ are derived as the $L_2([-N/2, N/2))$ inner-product between $s(\nu)$ and the shifted analysis functions $\tilde{\varphi}_p(\nu)$:

$$c[k] = \frac{1}{N}\int_{-N/2}^{N/2} s(\nu)\tilde{\varphi}_p^*\left(\frac{K}{N}\nu - k\right) d\nu. \tag{5}$$

Here, φ_p^* denotes the conjugate of the function φ_p. Note that both φ_p and $\tilde{\varphi}_p$ are N-periodic functions and need not be support limited. If φ_p and $\tilde{\varphi}_p$ are biorthogonal (i.e., $\langle\varphi_p((\nu)/(T) - k), \tilde{\varphi}_p((\nu)/(T) - l)\rangle = \delta_{k-l}$), then $x_{\text{app}}(\nu)$ is a projection of $s(\nu)$ onto the shift invariant space:

$$V_{\varphi_p} = \text{span}\left\{\varphi_p\left(\frac{K}{N}\nu - k\right); k = -K/2 \cdots K/2 - 1\right\}.$$

The synthesis equation in the periodic case (3) can also be written as (1), where the coefficients $c_p[k] = \sum_{n\in\mathbb{Z}} c[k+nK]$ is the K periodized version of $c[k]$. This is probably more familiar in the context of NUFFT; it denotes a simple interpolation using the finitely supported function ($\varphi \in L_2([-J/2, J/2])$) and periodic boundary conditions. This evaluation of a specific sample $s(\nu_m)$ requires J multiplications and additions.

B. Expression for the Approximation Error

The main factors that determine the quality of the shift-invariant approximation of the periodic signal $s(\nu)$ are as follows:

1) the interpolation function $\varphi(\nu) \in L_2([-J/2, J/2))$;
2) the sampling step $T = N/K$;
3) the K-periodic analysis function $\tilde{\varphi}_p(\nu)$.

A careful optimization of these factors is essential to minimize the approximation error, while keeping resources such as memory and computational complexity to a minimum. An exact expression for the approximation error is essential for a systematic computational optimization of the scheme.

We had derived the expression for the average error in approximating an arbitrary periodic function in a shift invariant (SI) space in [19]. The space V_{φ_p} is only integer shift variant; shifting the function $s(\nu)$ to $s(\nu - kT)$ (by noninteger multiples of the sampling step) affects the approximation error. The error is periodic with period $T = N/K$. The average mean square error[1] (averaged over all possible shifts) is shown in [19] as

$$\begin{aligned}\eta(s,\varphi,\tilde{\varphi},K) &= \frac{K}{N}\int_0^{N/K} \|s(\cdot-\tau) \\ &\quad - s_{\mathrm{app}}(\cdot-\tau)\|^2_{L_2([-N/2,N/2))} d\tau \\ &= \sum_{n=-\infty}^{\infty} |\hat{s}[n]|^2 E_{\varphi,\tilde{\varphi}}\left(\frac{2\pi n}{K}\right) \end{aligned} \tag{6}$$

where the error kernel $E_{\varphi,\tilde{\varphi}}(\omega)$ is given as

$$E_{\varphi,\tilde{\varphi}}(\omega) = \underbrace{1 - \frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}}_{E_{\min}(\omega)} + \underbrace{\hat{a}_\varphi(\omega)|\hat{\tilde{\varphi}}(\omega) - \hat{\varphi}_d(\omega)|^2}_{E_{\mathrm{res}}(\omega)}. \tag{7}$$

In (6), $\hat{s}[n]; n \in \mathbb{Z}$ are the Fourier series coefficients of $s(\nu)$, defined by

$$\hat{s}[n] = \frac{1}{N}\int_{-N/2}^{N/2} s(\nu) e^{\frac{-j2\pi\nu n}{N}} d\nu \tag{8}$$

and $\hat{\varphi}(\omega)$ is the Fourier transform of $\varphi(\nu)$:

$$\hat{\varphi}(\omega) = \int_{-\infty}^{\infty} \varphi(\nu) e^{-j\nu\omega} d\nu. \tag{9}$$

In (7), $\hat{a}_\varphi(\omega) = \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega + 2k\pi)|^2$ is the discrete Fourier transform of the autocorrelation sequence $a_\varphi(k) = \langle \varphi(\nu), \varphi(\nu - k)\rangle$. Moreover, $\hat{\varphi}_d(\omega) = \hat{\varphi}(\omega)/\hat{a}_\varphi(\omega)$ denotes the dual function of φ. Both $E_{\min}$ and E_{res} are positive terms.

[1]The variations in the error with respect to the shifts are shown to be small; the average error is a good indicator of the performance in most applications.

Moreover, $E_{\min}$ is only dependent on φ, while E_{res} is dependent of the analysis function $\tilde{\varphi}_p$; this term vanishes iff

$$\begin{aligned}\hat{\tilde{\varphi}}(2\pi n/K) &= \hat{\varphi}_d(2\pi n/K) \\ &= \frac{\hat{\varphi}(2\pi n/K)}{\sum_{k\in\mathbb{Z}} |\hat{\varphi}(2\pi n/K + 2k\pi)|^2};\ \forall n \in \mathbb{Z} \end{aligned} \tag{10}$$

With this specific choice of the analysis function, the error expression simplifies to

$$\eta_{\min}(s,\varphi,K) = \sum_{n=-\infty}^{\infty} |\hat{s}[n]|^2 \frac{\sum_{k\neq 0} |\hat{\varphi}(2\pi n/K + 2k\pi)|^2}{\sum_{k\in\mathbb{Z}} |\hat{\varphi}(2\pi n/K + 2k\pi)|^2}. \tag{11}$$

This is the minimum achievable mean square error for a specified signal $s(\nu)$, an interpolator φ and the sampling step K/N. Note that the metric (11) is independent of the analysis function $\tilde{\varphi}$. The approximation $x_{\mathrm{app}}(\nu)$, derived using of the optimal analysis function specified by (10), is the orthogonal projection of $s(\nu)$ onto V_{φ_p}.

C. NUFFT: Problem Statement

For simplicity, we restrict our attention to the 1-D NUFFT problem. We are given equally spaced samples $x[n]; n = -N/2 \ldots N/2 - 1$. The goal is to derive an efficient approximation of the discrete time Fourier transform (DTFT) of this sequence:

$$\hat{x}(\nu) = \sum_{-N/2}^{N/2-1} x[n] \exp\left(-\frac{j2\pi\nu n}{N}\right); \nu \in \mathbb{R}, \tag{12}$$

at the nonuniform frequency locations $\nu_m; m = 0..M-1$. Note that the conventions are slightly different from the previous section, where the discrete sequence $\hat{s}[n]$ indicates the Fourier series of the continuous domain function $s(\nu)$. In contrast, here we have $\hat{x}(\nu)$ as the discrete Fourier transform of the sequence $x[n]$. The direct evaluation of (12) is computationally expensive (requires $\mathcal{O}(MN)$ operations). Clearly, the continuous domain function $\hat{x}(\nu)$ is N-periodic. The frequency samples ν_m may thus be assumed to be in the range $[-N/2, N/2)$, without loss of generality. The sequence $x[n]$ may be obtained from the continuous domain function $\hat{x}(\nu)$ as

$$x[n] = \frac{1}{N}\int_{-N/2}^{N/2} \hat{x}(\nu) e^{\frac{j2\pi\nu n}{N}} d\nu. \tag{13}$$

D. Mathematical Formulation of the NUFFT Approximation

To reduce the computational cost in evaluating (12), the standard practice is to approximate it as an interpolation of the K-point uniform discrete Fourier transform (DFT) ($K \geq N$; K even) of $h[n]x[n]$:

$$c_p[k] = \sum_{n=-N/2}^{N/2-1} h[n]x[n] \exp\left(-\frac{j2\pi kn}{K}\right); k \in \mathbb{Z}. \tag{14}$$

Although this equation is valid for $k \in \mathbb{Z}$, it only needs to be evaluated for $k = K/2, \ldots, K/2-1$ because of the K-periodicity of $c_p[k]$. This summation is efficiently evaluated using FFT and has a computational complexity of $\mathcal{O}(K \log K)$. The weights $h[n]; n = -N/2, \ldots N/2-1$ are termed as scale-factors in the NUFFT literature. They are often chosen heuristically [12] or as a function of φ [13]. The NUFFT scheme thus approximates the exact DFT, specified by (12), as the interpolation of $c_p[k]$ using φ assuming periodic boundary conditions [12], [13], [25]

$$\hat{x}_{\text{app}}(\nu) = \sum_{k \in \mathbb{Z}} c_p[k] \varphi\left(\frac{K}{N}\nu - k\right). \tag{15}$$

Since $x[n]$ is support limited in the range $-N/2 \cdots N/2-1$, the ideal interpolator and scale-factors are $\varphi(\nu) = \text{sinc}(N\nu/K)/N$ and $h[n] = 1, \forall n$, respectively [13]. The NUFFT interpolation using these scale-factors and interpolator is exact. However, since the sinc function is not support limited, there is no computational gain in computing (15) over the direct evaluation. Hence, the general practice is to approximate the sinc function using support limited interpolators (support limited in the range $[-J/2, J/2]$) such as Kaiser–Bessel, Gaussian, or optimized functions [12], [13], [25]. The computational complexity in performing the Fourier interpolation at M nonuniform points is $\mathcal{O}(MJ)$. Nonuniform scale-factors are used here to compensate for the finite support.

E. Discretization of the Interpolator

The computation of (15) at arbitrary sampling locations ν_m requires the evaluation of φ at the locations $(K\nu/N - k); k \in \mathbb{Z}$. Since φ is support limited in the range $[-J/2, J/2]$, J evaluations of the interpolator at noninteger sampling locations are required to compute each Fourier sample $\hat{x}_{\text{app}}(\nu_m)$. The widely used interpolators have either complicated expressions or are obtained using optimization schemes. To prevent the evaluation of the interpolator from dominating the computational complexity, researchers often precompute the interpolators on a uniform grid, which is much finer than the grid on which the DFT is evaluated. Nonuniform samples of the interpolator are obtained by interpolating these lookup table entries using linear or nearest neighbor methods. We denote the oversampling factor by O and assume it to be an even integer. Thus, the interpolation function is modeled as (15)

$$\varphi(\nu) = \sum_{k=-JO/2+1}^{JO/2-1} q[k]\beta(O\nu - k) \tag{16}$$

where β is the B-spline function. The proposed theory is applicable for any discretization. We will specifically focus on linear and nearest neighbor interpolation, since they are the widely used schemes. The linear B-spline function is defined as

$$\beta^1(x) = \begin{cases} 1 - |x| & \text{if } |x| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

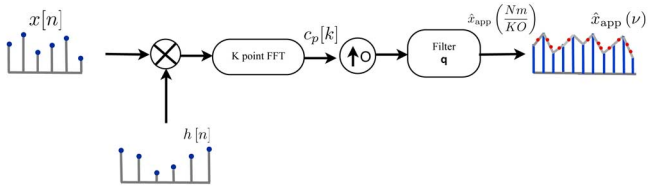

Fig. 1. Illustration of the NUFFT pipeline. The sequence $\mathbf{c}_p$ is evaluated as the K-point FFT of the input sequence $x[n]$, weighted by the scale-factors $h[n]$. $\mathbf{c}_p$ is over-sampled by O and interpolated to a fine uniform grid, using the discrete sequence $q[n]; n = -JO/2+1 \cdots JO/2-1$. These uniform samples $\hat{x}_{\text{app}}((Nm)/(KO))$ are then interpolated using linear or nearest neighbor schemes to obtain the nonuniform Fourier samples (denoted by the dots) $\hat{x}_{\text{app}}(\nu_m)$.

while the B-spline function of degree 0 is defined as

$$\beta^0(x) = \begin{cases} 1 & \text{if} |x| < 1/2 \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

The coefficients $q[k]$ in (16) are the samples of the specified function (Kaiser-Bessel/Gaussian) at the sampling locations k/O. In (16), we assumed that $\varphi(-J/2) = \varphi(J/2) = 0$. With the interpolation model (16), $\hat{x}_{\text{app}}(\nu)$ may also be thought of as the linear interpolation of the uniform samples $\hat{x}_{\text{app}}(Nm/KO); m \in \mathbb{Z}$:

$$\begin{aligned} &\hat{x}_{\text{app}}\left(\frac{Nm}{KO}\right) \\ &\quad = \sum_{k \in \mathbb{Z}} \sum_{n=-JO/2+1}^{JO/2-1} c_p[k]q[n]\beta(m - Ok - n) \\ &\quad = \sum_{k \in \mathbb{Z}} c_p[k]q[m - Ok]. \end{aligned} \tag{19}$$

We substituted (16) in (15) and used the relationship $\beta^0(m) = \beta^1(m) = \delta(m); m \in \mathbb{Z}$, where δ is the Kroneker delta function, to obtain the above equation. It implies that the uniform samples $\hat{x}_{\text{app}}((Nm)/(KO))$ are derived by up-sampling $c_p[k]$ by O and filtering it by the discrete filter $\mathbf{q}$. The entire NUFFT pipeline is illustrated in Fig. 1. In standard computational schemes, the last three steps (oversampling, filtering by $q[k]$ and interpolating using B-spline functions) are combined into a single complex operation for computational efficiency. We have chosen to expand these steps as in Fig. 1 to obtain a better understanding of the process and for the ease of analyzing the error involved in the discretization (see Section IV-C).

III. Optimal NUFFT Scale-Factors and Interpolator

In this section we introduce the framework to derive the optimal NUFFT scale-factors and interpolator. We derive the worst-case error in approximating a periodic signal in a shift invariant basis, which is then used to choose the optimal NUFFT parameters.

A. NUFFT as a Shift Invariant Approximation

We now show that the NUFFT approximation is essentially a shift invariant approximation of $\hat{x}(\nu)$, the original discrete Fourier transform of $x[n]$, in the shift invariant space V_{φ_p}.

Note that the NUFFT synthesis equation, specified by (15), is already in the same form as (1). Substituting $x[n]$ with the inverse Fourier transform of $\hat{x}(\nu)$ in (14), we obtain

$$c_p[k] = \frac{1}{N}\sum_{n=-N/2}^{N/2-1} \underbrace{\int_{-N/2}^{N/2} \hat{x}(\nu) e^{j\frac{2\pi\nu n}{N}} d\nu}_{x[n]} h[n] e^{-\frac{j2\pi kn}{K}}$$
$$= \frac{1}{N}\int_{-N/2}^{N/2} \hat{x}(\nu) \underbrace{\sum_{n=-N/2}^{N/2-1} h[n] e^{j\frac{2\pi}{K}\left(\frac{K\nu}{N}-k\right)n}}_{\tilde{\varphi}_p^*\left(\frac{K}{N}\nu-k\right)} d\nu. \quad (20)$$

Thus, the derivation of the coefficients in NUFFT can be interpreted as the $L_2([-N/2, N/2))$ inner-product between the exact continuous domain Fourier transform of $x[n]$, denoted by $\hat{x}(\nu)$, and the analysis functions $\tilde{\varphi}_p((K)/(N)\nu - k); k = -K/2, \ldots, K/2-1$. Thus, the NUFFT scale-factors are analogous to prefilters in shift invariant representations [26], [27]. Note that the equivalent analysis function $\tilde{\varphi}$ need not be support limited, even when φ is support limited.

B. Error Expression

Having shown that the NUFFT scheme is a shift invariant approximation, we now use the error metric introduced in Section II.B to analyze it. In contrast to the standard shift invariant setting, the NUFFT interpolation is performed in the Fourier domain; $\hat{x}(\nu)$ is the discrete Fourier transform of $x[n]$. Applying (6) to the representation of $\hat{x}(\nu)$, we obtain

$$\eta(x, \varphi, \tilde{\varphi}, K) = \sum_{n=-N/2}^{N/2-1} |x[n]|^2 E_{\varphi,\tilde{\varphi}}\left(-\frac{2\pi n}{K}\right). \quad (21)$$

We use the finite support of $x[n]$ to restrict the range of the summation. The negative sign in the index of $E_{\varphi,\tilde{\varphi}}$ is because of the difference in the definitions of $x[n]$ and $\hat{s}[n]$ [see (13) and (8)]. We now use this error expression to derive the optimal least-square scale-factors.

C. Optimum Least-Square Scale-Factors

From (10), it is easy to see that if the scale-factors $h[n]$ are chosen as

$$h[n] = \hat{\varphi}_d(-2\pi n/K)$$
$$= \frac{\hat{\varphi}(-2\pi n/K)}{\sum_{k\in\mathbb{Z}} |\hat{\varphi}(-2\pi n/K + 2k\pi)|^2}; \quad n = -N/2 \ldots N/2-1 \quad (22)$$

then $\hat{x}_{\text{app}}(\nu)$ is the orthogonal projection of $\hat{x}(\nu)$ onto the shift invariant space. This corresponds to the minimum achievable mean square error for any signal $x[n]$. The indexes of $\hat{\varphi}$ in (22) have negative signs because $h[n]$ is defined as the inverse discrete Fourier transform of φ_d in (20). We term this $h[n]$ as the optimal least-square (OLS) scale-factors. When the interpolator is a real valued function and its energy is concentrated in the Fourier domain (e.g., prolate spheroidal wave functions), the optimal scale-factors may be approximated (assuming $|\hat{\varphi}(2\pi n/K + 2k\pi)| \approx 0$ for $k \neq 0$ and $n = -N/2, \ldots, N/2-1$) as

$$h[n] \approx \frac{1}{|\hat{\varphi}(-2\pi n/K)|} = \frac{1}{|\hat{\varphi}(2\pi n/K)|}.$$

We also need to assume that $\hat{\varphi}(-\omega) = \hat{\varphi}(\omega)$, which is valid if φ is real. These approximate scale-factors are reported as the minimum aliasing error choice in [13]. The optimal LS scale-factors, specified by (22), provides the minimum possible error over all scale factors:

$$\eta_{\min}(x, \varphi, K) = \sum_{n=-N/2}^{N/2-1} |x[n]|^2 E_{\min}\left(-\frac{2\pi n}{K}\right). \quad (23)$$

Here, we used the symmetry of the error kernel.

D. Worst-Case Approximation Error

Note that (23) is dependent on the signal samples $x[n]$. We would like to have an expression that is only dependent on the interpolator. We now define the worst-case mean square error as the maximum of $\eta_{\min}(x, \varphi, K)$ over all possible sequences $x[n]$. Using Schwartz inequality, we obtain

$$\eta_{\max}^2(\varphi, K) = \arg\max_{x:\|x\|_{\ell_4}=1} \eta_{\min}(x, \varphi, K)$$
$$= \sum_{n=-N/2+1}^{N/2} \left|E_{\min}\left(\frac{2\pi n}{K}\right)\right|^2$$
$$= \sum_{n=-N/2+1}^{N/2} \frac{\left(\sum_{k\neq 0} |\hat{\varphi}\left(\frac{2\pi n}{K} + 2k\pi\right)|^2\right)^2}{\left(\sum_{k\in\mathbb{Z}} |\hat{\varphi}\left(\frac{2\pi n}{K} + 2k\pi\right)|^2\right)^2}. \quad (24)$$

Note that this is a tight bound for the approximation error. Sequences that satisfy

$$|x[n]|^2 = \frac{E_{\min}\left(-\frac{2\pi n}{K}\right)}{\sqrt{\sum_{l=-N/2+1}^{N/2} \left|E_{\min}\left(\frac{2\pi l}{K}\right)\right|^2}}; \quad n = -N/2, \ldots N/2-1$$

leads to the worst-case error. We show that (24) can be exactly and efficiently evaluated for any discretized interpolator in Section IV-A. Thus, it can be used to determine the optimal parameters of Kaiser–Bessel and Gaussian interpolators. Our main focus is to use this metric to derive the optimal least-square interpolator. Note that we derived the optimal scale-factors, before computing the worst-case error expression, only for simplicity. Since $E_{\min}$ and E_{res} are positive terms, the OLS scale-factors also leads to the minimum worst-case mean square error (over all possible scale-factors). The OLS scale-factors and the interpolators are optimal with respect to the worst-case mean square error criterion.

E. Optimized Least-Square Interpolator

Our goal is to derive a φ that is finitely supported in the range $[-J/2, J/2]$ and minimizes the worst-case LS error:

$$\varphi_{\text{OLS}} = \arg\min_{\varphi} \eta_{\max}^2(\varphi, K) \quad (25)$$

subject to $\|\varphi\|_{L_2[-J/2,J/2]} = 1$. Note that the functional $\eta_{\max}$ is nonquadratic. It is difficult to derive the continuous domain interpolator that minimizes the criterion. Moreover, obtaining the analytical expression of φ is not useful in a practical NUFFT setup, as discussed previously. Hence, we propose to numerically derive the optimal discretization of φ. We propose to use an iterative re-weighted minimization algorithm. Towards this end, we rewrite the criterion as a weighted quadratic criterion:

$$\eta^2_{\max}(\varphi,K) = \sum_{n=-N/2+1}^{N/2} \left(\sum_{k\neq 0} \left| \hat{\varphi}\left(\frac{2\pi n}{K} + 2k\pi\right)\right|^2 \right) \times \underbrace{\frac{\left(\sum_{k\neq 0} \left|\hat{\varphi}\left(\frac{2\pi n}{K} + 2k\pi\right)\right|^2\right)}{\left(\sum_{k\in\mathbb{Z}} \left|\hat{\varphi}\left(\frac{2\pi n}{K} + 2k\pi\right)\right|^2\right)^2}}_{f_\varphi[n]}. \quad (26)$$

We split the original criterion into two parts. We define the weights $f_\varphi[n]$ as

$$f_\varphi[n] = \begin{cases} \frac{\sum_{k\neq 0}\left|\hat{\varphi}\left(\frac{2\pi(n+kK)}{K}\right)\right|^2}{\left(\sum_k \left|\hat{\varphi}\left(\frac{2\pi(n+kK)}{K}\right)\right|^2\right)^2}; & n = -N/2+1,\ldots N/2 \\ 0 & \text{otherwise.} \end{cases} \quad (27)$$

At each iteration of the iterative algorithm, we assume the weights $f_\varphi[n]$ to be fixed and derive the φ that minimize the weighted quadratic criterion (26). The weights $f_\varphi[n]$ are then recomputed using (27). This process is repeated until the minimum of η^2 is reached. We do not make any theoretical claims on the convergence of the scheme to the global minimum. However, we observe that it converges to the same solution for all the initializations and parameter settings that we considered, unless $K = N$ (see Section V). The algorithm fails to converge when $K = N$. A more detailed description of the algorithm for discretized interpolators is given in Section IV-B.

F. Similarity to the Kaiser–Bessel Interpolator

If we set $f_\varphi[n] = 1; n = -N/2+1,\ldots,N/2$, then (26) can be rewritten as

$$\eta^2_{\max}(\varphi,K) = \underbrace{\sum_{n=-N/2+1}^{N/2} \sum_{k\in\mathbb{Z}} \left|\hat{\varphi}\left(\frac{2\pi(n+kK)}{K}\right)\right|^2}_{\approx 1} - \sum_{n=-N/2+1}^{N/2} \left|\hat{\varphi}\left(\frac{2\pi n}{K}\right)\right|^2 \quad (28)$$

$$\approx 1 - \sum_{n=-N/2+1}^{N/2} \left|\hat{\varphi}\left(\frac{2\pi n}{K}\right)\right|^2. \quad (29)$$

The first term in (28) is approximately equal to $\sum_{n\in\mathbb{Z}} |\hat{\varphi}((2\pi n)/(K))|^2 = \|\varphi\|_{L_2} = 1$. It is an equality when $K = N$; when $K > N$, the above approximation ignores some of the high frequency terms of φ, but may be valid for interpolators that are almost band-limited. The interpolator that minimizes (29) in-turn maximizes the in-band energy (energy in the range $-N/2\ldots N/2$). The well known zeroth order prolate spheroidal wave function (PSWF) $\phi_{J,\sigma}$ is a unit norm function ($\|\phi\|_{L_2} = 1$), which is finitely supported in the range $[-J/2, J/2]$ and have maximal energy in a specified range in the Fourier domain [28]

$$\int_{-\sigma}^{\sigma} |\hat{\phi}(\nu)|^2 d\nu. \quad (30)$$

Thus, the function obtained by the minimization of (29) may be seen as the discrete counterpart of the PSWF (29). The Kaiser–Bessel function was introduced as an approximation to the PSWF function [29]. Kaiser–Bessel functions are widely used in NUFFT and they give the best performance among other well known interpolators [13], [23].

IV. Numerical Evaluation of the Discretized Interpolator

In this section, we elucidate the iterative reweighted minimization scheme to derive the optimal discretization of the interpolator in more depth. We start by reducing the infinite summations in the worst-case error expression to finite sums to obtain an exactly computable metric.

A. Discretized Interpolator: Worst-Case Error Expression

To be consistent with the lookup table based evaluation of the NUFFT, we propose to derive φ in the space spanned by linear B-spline functions. Using the model specified by (16), the derivation of the OLS interpolator boils down to the derivation of the finite sequence $q[n]; n \in [-JO/2+1, JO/2-1]$. Computing the Fourier transform of (16), we obtain

$$\hat{\varphi}(\omega) = \frac{1}{O}\,\hat{q}\left(\frac{\omega}{O}\right)\hat{\beta}\left(\frac{\omega}{O}\right) \quad (31)$$

where $\hat{q}(\omega) = \sum_n q[n]\exp(-j\omega n)$ is the discrete time Fourier transform of $q[n]$. It is a 2π-periodic function, and hence $\hat{q}((x)/(O))$ is a $2O\pi$-periodic function. Using this model, we show in the Appendix that the infinite sums in the numerator and denominator of (24) are reduced to finite summations:

$$\sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega+2k\pi)|^2 = \sum_{k=-O/2}^{O/2} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}\right)\right|^2 \hat{a}_\beta\left(\frac{\omega+2k\pi}{O}\right) \quad (32)$$

$$\sum_{k\neq 0} |\hat{\varphi}(\omega+2k\pi)|^2 = \sum_{k=-O/2}^{O/2} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}\right)\right|^2 \hat{b}_\beta\left(\frac{\omega+2k\pi}{O}\right) \quad (33)$$

where $\omega = (2\pi n)/(K); n = -N/2,\ldots N/2$. Here, $\hat{a}_\beta$ and $\hat{b}_\beta$ are defined by (44) and (47). The functions $\hat{a}_\beta$ and $\hat{b}_\beta$ have exact expressions in terms of the autocorrelation of B-spline functions [30] (See the Appendix for details.) Using these relations, we

compute the weights $f_\varphi[n]$. Substituting from (33) in (26) and simplifying as shown in the Appendix, we obtain

$$\eta^2_{\max}(\varphi, K) = \sum_{n=-KO/2}^{KO/2} \left|\hat{q}\left(\frac{2\pi n}{KO}\right)\right|^2 g_q[n]. \quad (34)$$

This is a weighted norm in the Fourier domain, where the weights

$$g_q[n] = \hat{b}_\beta\left(\frac{2\pi n}{K}\right) \sum_{l\in\mathbb{Z}} f_\varphi[n - lK], \quad (35)$$

are assumed to be fixed at each iteration. We estimate the samples of the interpolator $q[l]; l = -JO/2 + 1, \ldots, JO/2 - 1$ such that this metric is minimized, subject to $||\varphi||_{L_2} = 1$.

B. Iterative Reweighted Minimization Algorithm

As discussed earlier, we propose an iterative algorithm that uses the following steps: Start with an initial sequence $\mathbf{q}_0$.

1) Set $i = 1$.
2) Using $\mathbf{q}_{i-1}$, derive the optimal weights $\mathbf{g}_{q_i}$ specified by (35).
3) Using the current weights $\mathbf{g}_{q_i}$, derive the optimal coefficient sequence $\mathbf{q}_{\text{opt}}$ that minimizes (34), subject to $||\varphi||_{L_2} = 1$. We show in the following section that the derivation of $\mathbf{q}_{\text{opt}}$ can be formulated as an eigendecomposition problem.
4) Derive $\mathbf{q}_i$ using the previous estimate $\mathbf{q}_{i-1}$ and $\mathbf{q}_{\text{opt}}$ as $\mathbf{q}_i = \alpha_i \mathbf{q}_{\text{opt}} + (1 - \alpha_i)\mathbf{q}_{i-1}; 0 \le \alpha_i \le 1$, where
$$\alpha_i = \arg\min_\alpha \left(\eta^2_{\max}(\alpha \mathbf{q}_{\text{opt}} + (1 - \alpha)\mathbf{q}_{i-1})\right). \quad (36)$$
5) Exit if $\alpha_i <$ TOLERANCE. Else, set $i = i + 1$.
6) Goto step 2.

Step 4 of the above algorithm is a one dimensional minimization procedure to ensure the monotonic convergence of the algorithm. The step-size α_i is determined using a simple linear search. We choose the initial sequence $\mathbf{q}_0$ as the samples of the optimized Kaiser-Bessel interpolator. We show in Section V that the final interpolator, to which the algorithm converges, is not dependent on the initialization. We now focus on step 3, where we derive the optimal $\mathbf{q}$ that minimizes (34) subject to $||\varphi||^2 = \mathbf{q}^T\mathbf{B}\mathbf{q} = 1$, for a specified $\mathbf{g}_\mathbf{q}$. This problem can be rewritten in the matrix form as

$$\mathbf{q}_{\text{OLS}} = \arg\min_{\mathbf{q}} \left(\mathbf{q}^T \underbrace{\mathbf{F}^T\mathbf{W}_g\mathbf{F}}_{\mathbf{A}}\mathbf{q}\right) \text{ subject to } \mathbf{q}^T\mathbf{B}\mathbf{q} = 1 \quad (37)$$

where $\mathbf{F}$ is the $KO \times JO - 2$ DFT matrix, $\mathbf{W}_g$ is the $KO \times KO$ diagonal matrix with diagonal entries as $\mathbf{g}_\mathbf{q}$. $\mathbf{B}$ is the autocorrelation matrix with entries $\mathbf{B}_{m,n} = \langle\beta(O\nu - m), \beta(O\nu - n)\rangle = a_\beta(m - n)/O$. $\mathbf{A}$ and $\mathbf{B}$ are matrices of dimension $(JO - 2) \times (JO - 2)$. Solving this constrained minimization problem using the Lagrange's multiplier method, we obtain

$$\mathbf{A}\mathbf{q}_{\text{opt}} = \lambda_{\min}\mathbf{B}\mathbf{q}_{\text{opt}} \quad (38)$$

where $\lambda_{\min}$ is the minimum generalized eigenvalue of the matrix pair $(\mathbf{A}, \mathbf{B})$ [31].

C. Error Due to the Discretization of the Interpolator

As discussed in Section II-E, we discretize the interpolator for its efficient evaluation within the NUFFT. The quality of the discretization depends on the oversampling factor. It is desirable that the minimum possible oversampling factor is chosen so as to minimize the size of the lookup table. At the same time, it is crucial that the interpolator is sampled at an adequately high rate so that the worst-case least-square error is not dominated by discretization. Due to these contradicting demands, it is highly desirable to quantify the error due to the discretization of the interpolator in the NUFFT approximation.

As seen in Section II-E, the approximation $\hat{x}_{\text{app}}(\nu)$ may also be interpreted as the linear interpolation of the sequence $\hat{x}_{\text{app}}((Nk)/(KO))$:

$$\hat{x}_{\text{app}}(\nu) = \sum_{k=-KO/2}^{KO/2-1} \hat{x}_{\text{app}}\left(\frac{Nk}{KO}\right)\beta(O\nu - k) \quad (39)$$

where $\hat{x}_{\text{app}}(Nk/KO)$ is obtained by convolving the up-sampled version of $c_p[k]$ with $\mathbf{q}$; $\mathbf{q}[n]$ are the uniform samples of the interpolator. In this two step process, there are two main error sources: a) the error in approximating the uniform samples of the DFT, (denoted by $\hat{x}(Nk/KO)$) by $\hat{x}_{\text{app}}(Nk/KO)$ and b) the error in approximating $\hat{x}(\nu)$ as the interpolation of the uniform samples $\hat{x}(Nk/KO)$. The first term is dependent on the specific choice of the sequence $\mathbf{q}$, while the second term is dependent on the B-spline interpolator β. To separate these error components, we rewrite (39) as

$$\hat{x}_{\text{app}}(\nu) = \sum_{k=-KO/2}^{KO/2-1} \left(\hat{x}(Nk/KO) - \underbrace{[\hat{x}(Nk/KO) - \hat{x}_{\text{app}}(Nk/KO)]}_{\hat{e}(Nk/KO)}\right)\beta(O\nu - k). \quad (40)$$

Using this relation, we obtain the total approximation error as

$$\begin{aligned}\mathcal{E}^2 &= ||\hat{x}(\nu) - \hat{x}_{\text{app}}(\nu)||^2 \\ &= \left\|\hat{x}(\nu) - \sum_{k=-KO/2}^{KO/2-1}[\hat{e}(Nk/KO) + \hat{x}(Nk/KO)]\beta(O\nu - k)\right\|^2 \\ &\le \left\|\sum_{k=-KO/2}^{KO/2-1}\hat{e}(Nk/KO)\beta(O\nu - k)\right\|^2 \\ &\quad + \underbrace{\left\|\hat{x}(\nu) - \sum_{k=-KO/2}^{KO/2-1}\hat{x}(Nk/KO)\beta(O\nu - k)\right\|^2}_{\mathcal{E}^2_{\text{discrete}}}.\end{aligned} \quad (41)$$

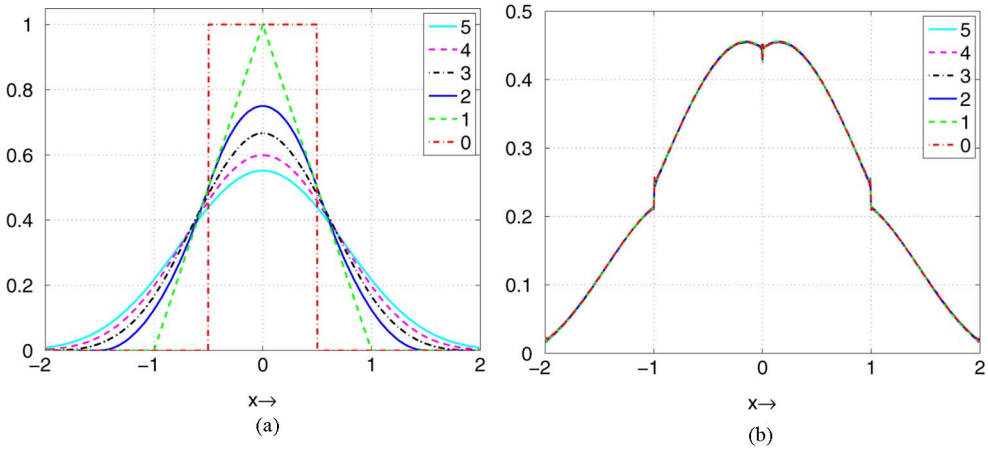

Fig. 2. Illustration of the convergence of the algorithm. The iterative reweighted algorithm was initialized with B-spline functions of degree 0 to 5, indicated by the curves in a. We derived the OLS interpolator using the iterative reweighted minimization algorithm, assuming $K = 132, N = 128, O = 100, J = 4$. Note that the final solution is the same, irrespective of the initialization. The worst-case errors are also the same, up to numerical precision. (a) Initializations. (b) Final solution.

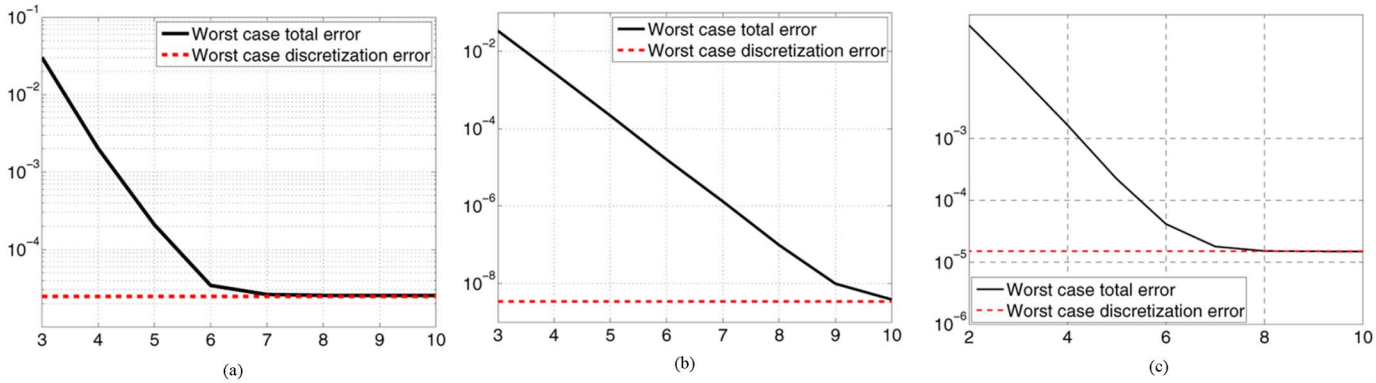

Fig. 3. Effect of discretization on the worst-case error. We considered $K = 140$ and $N = 128$. (a) indicates the decay of the worst-case OLS errors as a function of J, when $O = 10$ and linear interpolation is used. It is seen that the discretization error dominates the total error for interpolators with $J \geq 6$. (b) shows the worst-case error curve for $O = 100$ and linear interpolation. The performance of the NUFFT is significantly improved in this case. (c) indicates the decay for the same settings, but nearest neighbor interpolation. It is seen that the error saturates to a slightly lower value that the linear interpolation case with $O = 10$. (a) $O = 10$; Linear interpolation. (b) $O = 100$ Linear interpolation (c) $O = 100$; nearest neighbor interpolation.

Here, we used the triangle inequality of the L_2 norm. Thus, the first term in (41) is dependent on the specific choice of $\mathbf{q}$, while the second term $\mathcal{E}^2_{\text{discrete}}$ is independent of $\mathbf{q}$. It is dependent on the sampling interval KO/N, the discrete sequence $x[n]$, and the B-spline function β. Using the same argument as in (24), we upper bound of this term as

$$\mathcal{E}^2_{\text{discrete}}(O) \leq \sum_{n=-N/2}^{N/2} \frac{\left(\sum_{k\neq 0}\left|\hat{\beta}\left(\frac{2\pi n}{KO}+2k\pi\right)\right|^2\right)^2}{\left(\sum_{k\in\mathbb{Z}}\left|\hat{\beta}\left(\frac{2\pi n}{KO}+2k\pi\right)\right|^2\right)^2}. \tag{42}$$

We compare this term with the total worst-case NUFFT error specified by (24) to see if it is the dominant component. As a thumb rule, we seek to keep the worst-case discretization error around 10 times smaller that (24) by appropriately choosing the oversampling factor O and the degree of the B-spline interpolator.

D. Utilizing Symmetry to Reduce Computational Cost

The evaluation of step three of the iterative reweighted algorithm (described in Section IV-B) involves the eigendecomposition of the matrix pair $(\mathbf{A}, \mathbf{B})$, specified by (38). The computational complexity and the numerical stability of the eigen decomposition grows significantly with the sizes of $\mathbf{A}$ and $\mathbf{B}$. The number of unknowns (length of $\mathbf{q}$) can be reduced by a factor of two by assuming the interpolation functions to be symmetric. The matrices $\mathbf{A}$ and $\mathbf{B}$ can also be modified to account for the symmetry, thus reducing their size by a factor of 2 in either dimensions. This approach significantly reduces the computation time (often by a factor of 10 or more). The decrease in computational complexity becomes even more significant for high values of J and O. We demonstrate that there is no performance loss in using the symmetry constraint in the next section.

V. Results

In this section, we analyze the convergence of the iterative reweighted algorithm, the effect of discretization and the use of symmetry constraint. We also compare the OLS NUFFT scheme

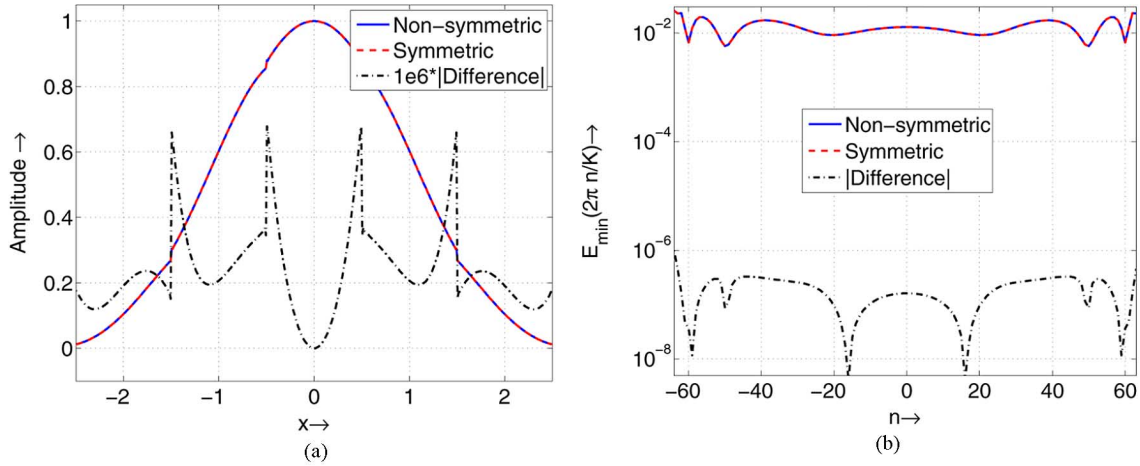

Fig. 4. Comparison of symmetric and nonsymmetric interpolators at $K = 132 N = 128 J = 5$ and $O = 100$. Both the algorithms were initialized with the optimized Kaiser–Bessel function. Note that the shape of the optimal interpolators and the error kernels are overlapping. The scaled (by 1e6) differences between the functions and the absolute differences between the kernels are also shown. (a) Interpolators. (b) Error kernels.

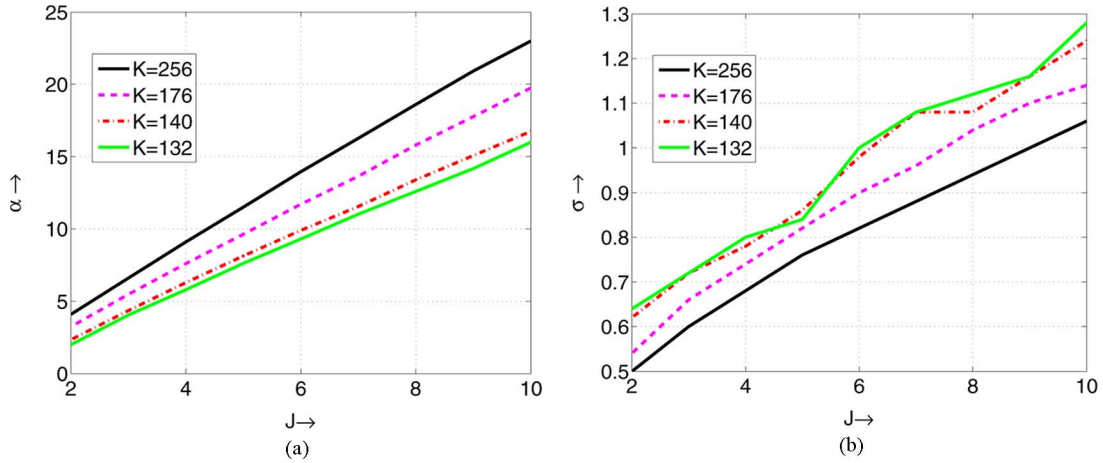

Fig. 5. Optimal least-square parameters of the Kaiser–Bessel and Gaussian functions. These parameters are assumed in all the comparisons considered in this section. (a) Optimal KB parameters. (b) Optimal Gaussian parameters.

with standard approximation methods using theoretical metrics as well as numerical methods.

A. Analysis of the Iterative Reweighted Algorithm

We first demonstrate the convergence of the algorithm on a specific example $(K = 132, N = 128, O = 100, J = 4)$. We consider a wide range of input initializations (ranging from B-spline of order 0 to order 5). It is seen from Fig. 2 that the final solution is the same, irrespective of the initialization. We performed similar comparisons for a wide range of parameter sets (K, N, and O) and verified that the algorithm converges to the same function in all the cases, except when $K = N$. In this case, the iterative reweighted minimization scheme fails to converge.

We study the effect of the discretization of the interpolators on NUFFT performance in Fig. 3. We considered $K = 140, N = 128$ and varied J from 3 to 10. The worst-case total errors of the OLS interpolators and the corresponding worst-case discretization errors are plotted for (a) $O = 10$ and linear interpolation, (b) $O = 100$ and linear interpolation, and (c) $O = 100$ and nearest neighbor interpolation. It is seen from Fig. 3(a) that for $O = 10$, the worst-case total error saturates to the worst-case discretization error, when $J > 6$. The comparison of the error metrics specified by (42) and (24) enables us to determine whether the performance of the NUFFT is limited by the discretization. Note that (42) is a good indicator of the influence of discretization on the total error. Note from Fig. 3(b) that choosing $O = 100$ significantly decreased the errors over $O = 10$. However, some saturation effects can be seen as the worst-case total error curve approaches the worst-case discretization error. It is seen in that even better results are obtained by further increasing O to 170 in Fig. 7. The nearest neighbor interpolation using $O = 100$ performs only slightly better than the linear interpolation with $O = 10$ [see Fig. 3(c)]. Hence, we will assume linear interpolation for all the comparisons.

We now show that there is practically no difference in the performance between interpolators optimized with and without the symmetry constraint in Fig. 4. In contrast, the time taken to evaluate the functions on a Macintosh 2.33-GHz Intel Core2Duo processor is 14 and 143 s, respectively. The drastic reduction in complexity is due to the decreased size of $\mathbf{A}$ and $\mathbf{B}$ matrices, leading to eigen decomposition of a smaller system. We have

performed similar studies for a range of parameter settings and verified that the use of symmetry constraint does not decrease the performance. The performance improvement in using the symmetry constraint increases with J, O and K. Hence, we restrict our attention to symmetric interpolators in the rest of the paper.

B. Comparison With Standard Interpolators

We compare the performance of the OLS function with Kaiser–Bessel, Gaussian and min–max interpolators in this subsection. To perform objective comparisons, we optimize the Kaiser–Bessel (KB) and Gaussian interpolators with respect to the mean square criterion (see Fig. 5). We assume the KB order to be zero and determine the optimal value for α. It is seen that the optimal α varies almost linearly with J from Fig. 5(a), while the slope is different for different values of K. When $K = 2N$, we find that the slope is approximately 2.30, close to the optimal value determined by Fessler *et al.* [13]. The parameter σ of the Gaussian function $(\exp(-x^2/\sigma^2)/2)$ corresponding to the different parameter settings are also shown in Fig. 5(b). We also use the OLS scale-factors (22) for these interpolators.

For the min–max interpolators, we use the scale-factors with the highest order that is reported in [13], when $K = 2N$. For $K < 2N$, we derive the scale factors by performing a least-square fit of the parametric model $(\beta = 1; L = 22)$ to the optimal Kaiser–Bessel scale-factors (see [13] for details). The default of $L = 13$ gives interpolation errors that are comparable to uniform scale factors. However, increasing L beyond 22 results in poor fit to the KB scale factors due to the bad conditioning of the system matrix.

The interpolators and their error kernels at $K = 132$ and $K = 256$ are shown in Fig. 6. It is seen that the subtle variations in the shape of the interpolators (Kaiser–Bessel, Gaussian and the OLS functions) in Fig. 6(a) lead to significant discrepancies in the error kernels. In contrast to the OLS function, the error kernels of the standard interpolators are significantly elevated close to the edge of the signal as shown in Fig. 6(b), thus resulting in higher worst-case errors. By spreading the error to all spatial locations, the OLS interpolator significantly reduces the worst-case mean square error. The OLS interpolator gave lower errors at most spatial locations when compared with the other functions, even when $K = 256$. However, the performance improvement in this case is not as drastic as in (a-b). It is seen from Fig. 6(e) and (f) that the error kernels of the min–max interpolator, derived with the assumption of third order scale-factors, are comparable to the OLS function.

In Fig. 7, we compare the different interpolator families based on the worst-case mean square error (24) and the min–max error (derived in [13]). The comparisons were performed at $K = 132$, $K = 140$ and $K = 256$, respectively. Fig. 7(a) and (d) denotes the error curves at K $= 132$ and oversampling factor $= 100$. The use of a high oversampling factor ensures that this term is not the significant contributor. Similar comparisons are shown for $K = 140$ and $K = 256$ in Fig. 7(b) and (e) and Fig. 7(c) and (f), respectively. Note that although we minimized the worst-case mean square errors, the proposed NUFFT scheme performs very well with respect

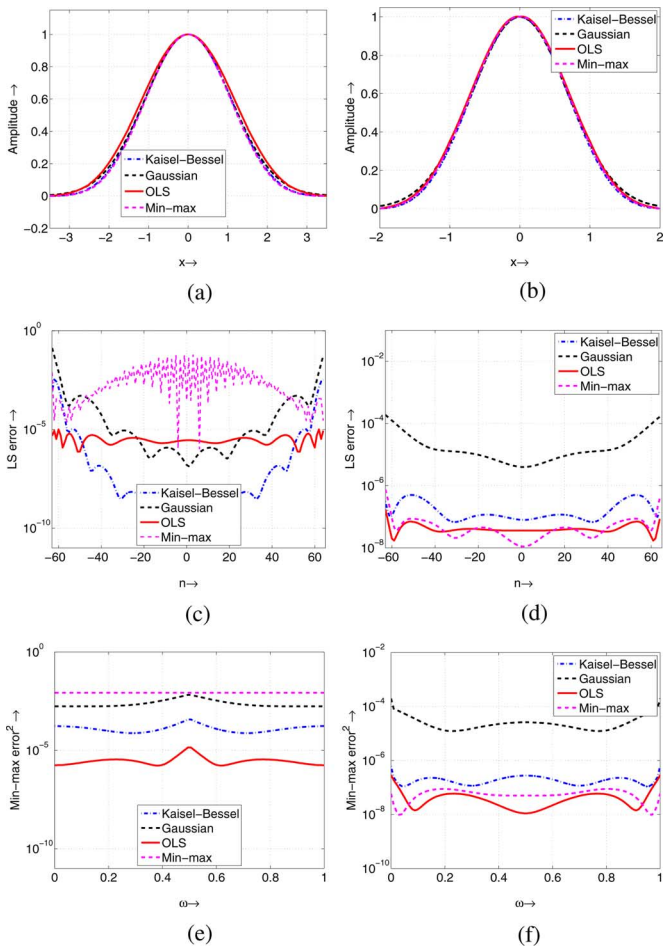

Fig. 6. Comparison of the OLS interpolator with classical schemes. (a), (c), and (e) compares the interpolators at $K = 132N = 128, O = 100$ and $J = 7$. (a) indicates the shape of the interpolators (c). Comparison based on the mean square error kernel. (e). Comparison based on the min–max error criterion derived in [13]. (b), (d), and (f): Comparison of the interpolators for $K = 256N = 128J = 4$ and $O = 190$. (d) indicates shape of the interpolators while (e) and (f) shows the mean square error kernel and the min–max error kernels. (a) Interpolator: $K = 132; J = 7$. (b) Interpolator: $K = 256; J = 4$. (c) MS error kernel: $K = 132; J = 7$. (d) MS error kernel: $K = 256; J = 4$. (e) MM error kernel: $K = 132; J = 7$. (f) MM error kernel: $K = 256; J = 4$.

to both the error metrics at almost all parameter settings. The min–max estimator provides lower min–max errors for lower values of J. However as the length of the interpolator increases, its performance deteriorates. This is probably due to the insufficient accuracy of the least-square fit to the optimal Kaiser-Bessel scale factors. As explained previously, the use $L = 22$ instead of the default value $L = 13$ improves the performance min–max interpolators. In the $K = 256$ case, we used the best scale-factors that were available for the min–max interpolator. Third order scale-factors were used for $J = 4$ and $J = 6$, while only second order scale-factors were available for the rest (this explains the oscillatory nature of the mean square error). Note that the min–max interpolator gives comparable errors with the OLS function for $K = 256$ and lower values of J. However, its performance saturates for higher values of J, probably because of the limited number of scale-factor parameters that can be derived in the min–max setting. It is seen from Fig. 7(f) that the min–max interpolator performs better than the OLS function with respect to the maximum worst-case

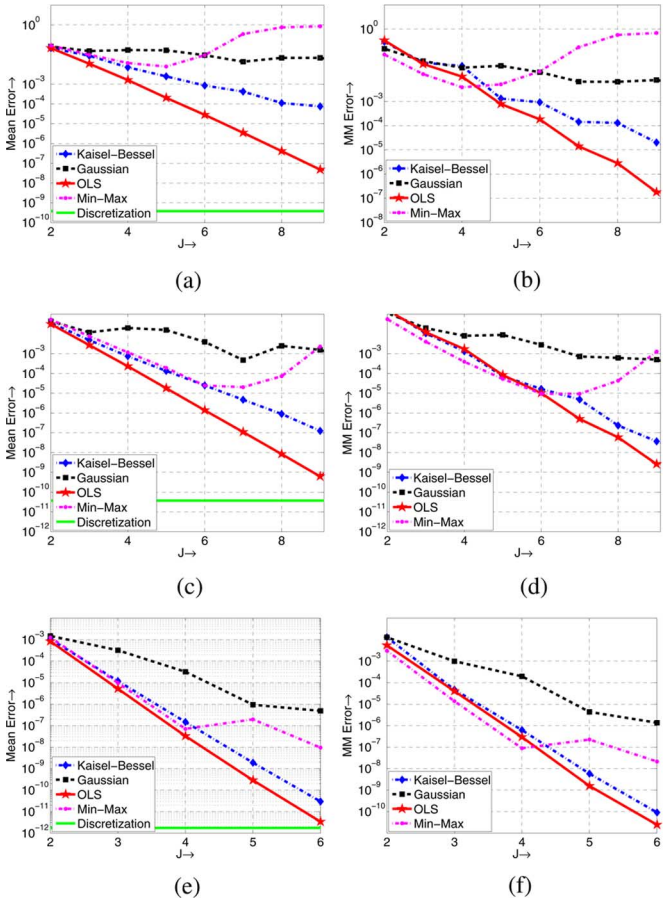

Fig. 7. Theoretical comparison of the different interpolator families. (a) and (b) indicate the decay of the worst-case mean square and min–max errors respectively for $\mathrm{K} = 132$ and oversampling $\mathrm{factor} = 100$. The worst-case discretization error is displayed in magenta. (c), (d): Comparison of errors for $K = 140$ and $O = 170$. As mentioned previously, we used uniform scale-factors for the min–max interpolators for ($K = 132$ and $K = 140$), while optimized min–max scale-factors from [13] were used for $K = 2N$. (e), (f) Comparison of different kernels at $\mathrm{K} = 256, \mathrm{N} = 128$ and $O = 190$. (a) $\mathrm{K} = 132$. (b) $\mathrm{K} = 132$. (c) $\mathrm{K} = 140$. (d) $\mathrm{K} = 140$. (e) $\mathrm{K} = 256$. (f) $\mathrm{K} = 256$.

error metric, when $J < 4$. However, note from Fig. 6(f) that the min–max function gives higher errors than the OLS function at most frequency locations for the $J = 4$ case (except close to the signal boundary).

It is seen from the comparisons that the proposed OLS NUFFT scheme significantly outperforms its closest competitor: the Kaiser–Bessel interpolator. The performance improvement is more significant for small values of K. For example, the use of the OLS interpolator provides approximately a factor of 3×10^3 decrease in the mean-square error at $J = 9; K = 132$. The interpolator settings ($\mathrm{J} = 9; \mathrm{K} = 132$) provides a worst-case error, which is comparable to that obtained with $J = 5$ at $K = 256$. Since lower value of K/N implies NUFFT algorithms with lower memory demands, these cases are of foremost interest in practical applications.

C. OLS Interpolators for Different Parameter Settings

We now analyze the OLS interpolators in more detail. In Fig. 8, we plot them for different values of J and K/N. The interpolators are normalized so that their integral is one. Note from Fig. 8(c) that the functions are discontinuous at the boundaries and at the origin for lower values of J. As J increases,

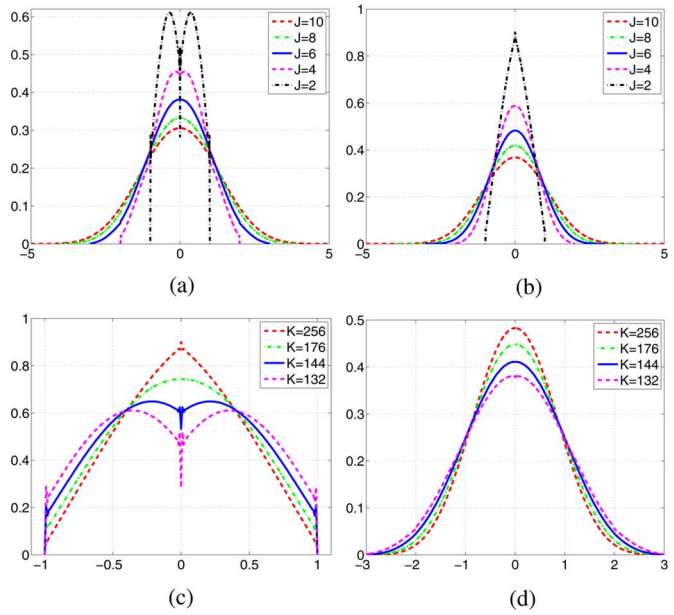

Fig. 8. Shape of the interpolators at different values of J and K. The figures in the top row, (a) and (b) show the interpolators for different values of J, corresponding to $K = 132$ and $K = 256$, respectively. The bottom row shows the interpolators at different values of K for $J = 2$ and $J = 6$ respectively. (a) $\mathrm{k} = 132$. (b) $\mathrm{k} = 256$. (c) $\mathrm{J} = 2$. (d) $\mathrm{J} = 6$.

they become smoother [see from Fig. 8(d)]. It is seen that the $J = 2$ interpolator changes its shape significantly with K; when $K = 256$, it approximates the linear B-spline function.

D. Numerical Simulations

We now compare the numerical performance of the interpolators in a simple one dimensional experiment. We considered the center line [shown in Fig. 9(a)] of the standard 128×128 Shepp–Logan phantom. The Fourier transform of this phantom is evaluated at 10 000 uniformly distributed random points (in the range $[-64, 63]$) in the Fourier domain. We consider the exact Fourier transform of the sequence at these points as the ground truth. The NUFFT approximations using the OLS and the standard interpolators optimized with respect to the worst-case mean square error are compared to the ground truth. We plot the errors as a function of J and K/N in Fig. 9. Note that the NUFFT interpolation using the OLS interpolator provides the best performance, consistent with theoretical worst-case predictions. Note that the error curves follow the same pattern as predicted by theory. The NUFFT approximation at $K = 140$ and $J = 10$ provides almost the same mean-squares error as $K = 256$ and $J = 10$.

VI. Discussion and Conclusion

The main focus of this paper was to derive a memory efficient approximation for the nonuniform Fourier transform of a discrete sequence. We derived an exact and computable expression for the worst-case mean square error in approximating the exact Fourier transform using the nonuniform fast Fourier transform (NUFFT) method. This metric was used to derive the optimal NUFFT interpolator and scale-factors, thus resulting in an algorithm with lower approximation errors. We also quantified the error in discretizing the interpolator. This measure ensured that a lookup table of sufficient size is used so that the dis-

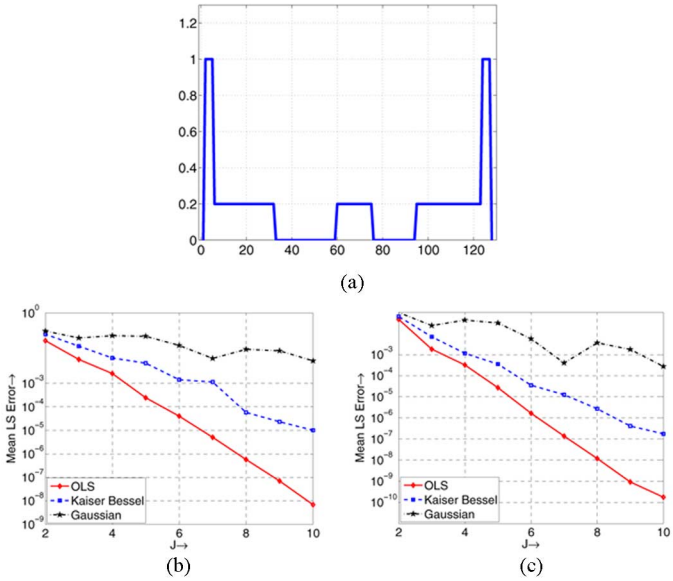

Fig. 9. Numerical comparison of the different interpolators using the center line of the Shepp–Logan phantom shown in (a). The errors in the comparisons (between the exact evaluations and the NUFFT approximations) are plotted for different values of J and K. (b) $\mathrm{K} = 132$ (c) $\mathrm{K} = 140$.

cretization error is not the dominant component. We compared theoretical and numerical performance of the proposed NUFFT scheme with standard methods. The comparisons clearly indicated that the proposed method significantly reduces the errors, especially when the oversampling factor is small. Thus, the proposed NUFFT scheme can result in iterative signal reconstruction schemes with much reduced memory requirements.

We focused on the derivation of one-dimensional interpolator in this paper. A simple strategy to extend this scheme to multiple dimensions is to consider the tensor product of 1-D interpolators [13]. This interpolator will inherit the error minimizing properties of the 1-D scheme.

The computationally demanding components of the NUFFT algorithm are (a) the evaluation of the uniform FFT and (b) the interpolation in the Fourier domain. Increasing the length of the uniform FFT augments the computational complexity faster than linear, while the cost of the interpolation increases linearly with the support of the interpolator. From our numerical studies, we found the computational complexity of the algorithm for $K = 140, J = 10$ and $K = 256, J = 5$ to be roughly the same. Note that these parameter choices gave approximations with almost the same errors. Moreover, the Fourier interpolations may be greatly accelerated using GPU implementations [32]. Of these two cases, note that the $K = 140, J = 10$ OLS NUFFT provides a factor of 1.8^d decrease in the required memory for d-dimensional data-sets over the $K = 256, J = 5$ OLS NUFFT scheme.

APPENDIX

A. Derivation of (32) and (33)

Using the Fourier transform of the model for φ, specified (16), the denominator of (24) becomes

$$\sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega+2k\pi)|^2$$

$$= \frac{1}{O^2}\sum_{k\in\mathbb{Z}} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}\right)\right|^2 \left|\hat{\beta}\left(\frac{\omega+2k\pi}{O}\right)\right|^2$$

$$= \frac{1}{O^2}\sum_{l\in\mathbb{Z}}\sum_{k-O/2}^{O/2} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}+2l\pi\right)\right|^2 \times \left|\hat{\beta}\left(\frac{\omega+2k\pi}{O}+2l\pi\right)\right|^2$$

$$= \sum_{k=-O/2}^{O/2} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}\right)\right|^2 \hat{a}_\beta\left(\frac{\omega+2k\pi}{O}\right). \quad (43)$$

Here, we used the 2π periodicity of $\hat{q}$. Here, a_β, specified by

$$\hat{a}_\beta(\omega) = \frac{1}{O^2}\sum_{l=-\infty}^{\infty} |\hat{\beta}(\omega+2l\pi)|^2 \quad (44)$$

is the DFT of the sequence $\langle \beta^n(x), \beta^n(x-l)\rangle = \beta^{(2n+1)}(l)$. B-spline functions have analytical expressions and hence can be evaluated analytically at any specified location [30]. The kernel need to be evaluated for $|\omega| \le \pi$. We rewrite the numerator as the difference between the denominator term, derived above, and $|\varphi(\omega)|^2$:

$$\sum_{k\neq 0} |\hat{\varphi}(\omega+2k\pi)|^2$$

$$= \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega+2k\pi)|^2 - |\hat{\varphi}(\omega)|^2$$

$$= \sum_{k=-O/2}^{O/2} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}\right)\right|^2 \hat{a}_\beta\left(\frac{\omega+2k\pi}{O}\right) - \frac{1}{O^2}\left(\left|\hat{q}\left(\frac{\omega}{O}\right)\right|^2 \left|\hat{\beta}\left(\frac{\omega}{O}\right)\right|^2\right). \quad (45)$$

For $|\omega| \le \pi$, this term can be rewritten as

$$\sum_{k\neq 0} |\hat{\varphi}(\omega+2k\pi)|^2 = \sum_{k=-O/2}^{O/2} \left|\hat{q}\left(\frac{\omega+2k\pi}{O}\right)\right|^2 \hat{b}_\beta\left(\frac{\omega+2k\pi}{O}\right). \quad (46)$$

where

$$\hat{b}_\beta(\omega) = \begin{cases} \hat{a}_\beta(\omega) - \frac{|\hat{\beta}(\omega)|^2}{O^2}; & |\omega| < \pi/O \\ \hat{a}_\beta(\omega); & \text{otherwise.} \end{cases} \quad (47)$$

B. Derivation of (34)

Substituting from (46) in (26), we obtain

$$\eta^2_{\max}(\varphi, K) = \sum_{n=-N/2+1}^{N/2} f_\varphi[n] \sum_{k=-O/2}^{O/2} \left|\hat{q}\left(\frac{2\pi(n+kK)}{KO}\right)\right|^2 \times \hat{b}_\beta\left(\frac{2\pi(n+kK)}{KO}\right)$$

$$= \sum_{k=-O/2}^{O/2} \sum_{n=kK-N/2+1}^{kK+N/2} \left|\hat{q}\left(\frac{2\pi n}{KO}\right)\right|^2$$

$$\times \hat{b}_\beta\left(\frac{2\pi n}{KO}\right) f_\varphi[n-kK]. \quad (48)$$

In the last equation, we used a change of variables $n+kK \to n$ to simplify the expression. Since f_φ is finitely supported [see (27)], we rewrite the term $f_\varphi[n-kK]$ as $\sum_{l\in\mathbb{Z}} f_\varphi[n-lK]$, thus making it independent of k. Moreover, we combine the summations (exploiting the structure of the support of f_φ) to obtain

$$\eta^2_{\max}(\varphi,K) = \sum_{n=-KO/2}^{KO/2} \left|\hat{q}\left(\frac{2\pi n}{KO}\right)\right|^2 \times \underbrace{\hat{b}_\beta\left(\frac{2\pi n}{KO}\right)\sum_{l\in\mathbb{Z}} f_\varphi[n-lK]}_{g_q[n]}. \quad (49)$$

ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers, whose constructive comments have ameliorated the paper.

REFERENCES

[1] S. Matej, J. Fessler, and I. Kazantsev, "Iterative tomographic image reconstruction using Fourier-based forward and back-projectors," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 401–412, Apr. 2004.

[2] Y. O'Connor and J. Fessler, "Fourier-based forward and back-projectors in iterative fan-beam tomographic image reconstruction," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 582–589, May 2006.

[3] K. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magn. Resonance Med.*, vol. 46, no. 4, pp. 638–51, Oct. 2001.

[4] H. Eggers, T. Knopp, and D. Potts, "Field inhomogeneity correction based on gridding reconstruction for magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 26, no. 3, pp. 374–384, Mar. 2007.

[5] B. Sutton, D. Noll, and J. Fessler, "Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 178–188, Feb. 2003.

[6] D. C. Munson, Jr. and J. L. C. Sanz, "Image reconstruction from frequency-offset Fourier data," *Proc. IEEE*, vol. 72, no. 6, pp. 661–669, Jun. 1984.

[7] T. Strohmer, "Computationally attractive reconstruction of bandlimited images from irregular samples," *IEEE Trans. Image Process.*, vol. 6, no. 4, pp. 540–548, Apr. 1997.

[8] H. Feichtinger and T. Strohmer, "Efficient numerical methods in non-uniform sampling theory," *Numer. Math.*, vol. 69, no. 4, pp. 423–440, Jan. 1995.

[9] S. Bagchi and S. K. Mitra, *The Nonuniform Discrete Fourier Transform and Its Applications in Signal Processing*, ser. Springer Int. Series in Engineering and Computer Science. New York: Springer, Nov. 1998.

[10] A. Makur and S. Mitra, "Warped discrete-Fourier transform: Theory and applications," *IEEE Trans. Circuits Syst. I*, vol. 48, no. 9, pp. 1086–1093, Sep. 2001.

[11] E. Candes and D. Donoho, "Ridgelets: A key to higher-dimensional intermittency?," *R. Soc. Lond. Philos. Trans. A, Math. Phys. Eng. Sci.*, vol. 357, no. 1760, pp. 2495–2509, 1999.

[12] N. Nguyen and Q. Liu, "The regular Fourier matrices and nonuniform fast Fourier transforms," *SIAM J. Sci. Comput.*, vol. 21, no. 1, pp. 283–293, 1999.

[13] J. Fessler and B. Sutton, "Nonuniform fast Fourier transforms using min–max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.

[14] A. Dutt and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SIAM J. Sci. Comput.*, vol. 14, no. 6, pp. 1368–1393, 1993.

[15] D. Potts, G. Steidl, and M. Tasche, *Fast Fourier Transforms for Nonequispaced Data: A Tutorial*, J. J. Benedetto and P. J. S. G. Ferreira, Eds. Boston, MA: Birkhäuser, 2001, pp. 247–270, "Fast Fourier transforms for nonequispaced data: A tutorial," *Modern Sampling Theory: Mathematics and Applications*, Jan. 2000.

[16] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, Jan. 1996, p. 490.

[17] M. Unser, "Sampling-50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.

[18] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques. I. Interpolators and projectors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2783–2795, Oct. 1999.

[19] M. Jacob, T. Blu, and M. Unser, "Sampling of periodic signals: A quantitative error analysis," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1153–1159, May 2002.

[20] A. Nieslony and G. Steidl, "Approximate factorizations of fourier matrices with nonequispaced knots," *Linear Algebra Appl.*, vol. 366, pp. 337–351, 2003.

[21] G. Steidl, "A note on fast Fourier transforms for nonequispaced grids," *Adv. Comput. Math.*, vol. 9, no. 3–4, pp. 337–352, 1998.

[22] J. Jackson, C. Meyer, D. Nishimura, and A. Macovski, "Selection of a convolution function for fourier inversion using gridding [computerised tomography application]," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 473–478, 1991.

[23] P. Beatty, D. Nishimura, and J. Pauly, "Rapid gridding reconstruction with a minimal oversampling ratio," *IEEE Trans. Med. Imag.*, vol. 24, no. 6, pp. 799–808, Jun. 2005.

[24] G. Chuang and C. Kuo, "Wavelet descriptor of planar curves: Theory and applications," *IEEE Trans. Image Process.*, vol. 5, no. 1, pp. 56–70, Jan. 1996.

[25] T. Strohmer, "Numerical analysis of the non-uniform sampling problem," *J. Comput. Appl. Math. (Special Issue on Numerical Analysis 2000, Vol. II: Interpolation and Extrapolation)*, vol. 122, no. 1–2, pp. 297–316, 2000.

[26] M. Unser and I. Daubechies, "On the approximation power of convolution-based least squares versus interpolation," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1697–1711, Jul. 1997.

[27] C. Lee, M. Eden, and M. Unser, "High-quality image resizing using oblique projection operators," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 679–692, May 1998.

[28] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V-the discrete case," *Bell Syst. Tech. J.*, vol. 57, pp. 1371–1430, May–Jun. 1978.

[29] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice–Hall, Jan. 1989.

[30] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[31] G. Strang, *Introduction to Linear Algebra*. Cambridge, MA: Wellesley-Cambridge, Apr. 2003.

[32] T. SØrensen, T. Schaeffter, K. Noe, and M. Hansen, "Accelerating the nonequispaced fast Fourier transform on commodity graphics hardware," *IEEE Trans. Med. Imag.*, vol. 27, no. 4, pp. 538–547, Apr. 2008.

Mathews Jacob (M'99) received the M.E. degree in signal processing from the Indian Institute of Science, Bangalore, in 1999 and the Ph.D. degree from the Biomedical Imaging Group at the Swiss Federal Institute of Technology , Lausanne, Switzerland, in 2003.

He was a Beckman Postdoctoral Fellow at the University of Illinois at Urbana Champaign between 2003 and 2006. Since 2007, he has been an Assistant Professor at the Department of Biomedical Engineering and Imaging Sciences at the University of Rochester, NY. His research interests include image reconstruction, image analysis and quantification in the context of a range of modalities including magnetic resonance imaging, near-infrared spectroscopic imaging and electron microscopy.

Dr. Jacob is currently the Associate Editor of the IEEE TRANSACTIONS ON MEDICAL IMAGING.