OPTIMIZED NON-UNIFORM FAST FOURIER TRANSFORM (NUFFT) FOR ITERATIVE TOMOGRAPHIC RECONSTRUCTION

Mathews Jacob

Department of Biomedical Engineering
University of Rochester

ABSTRACT

The main focus of this paper is the efficient approximation of the non-uniform Fourier transform (NUFFT). We reformulate the standard NUFFT approximation as a projection of the exact discrete Fourier transform onto a shift-invariant space. This reformulation enables the use of sophisticated tools, developed in the context of shift-invariant representations, to analyze the performance of the approximation. Using these techniques, we derive the optimal scale factors for a specified interpolator. Assuming these scale factors, we develop a worst-case error criterion that is only dependent on the interpolating function. We propose an iterative re-weighted optimization algorithm to derive the optimized least square (OLS) interpolator. This interpolator significantly reduces the approximation error in comparison to the standard methods. The improved performance of this scheme, for low oversampling factors, could lead to a memory efficient algorithm for non-Cartesian Fourier inversion.

Index Terms— sampling, interpolation, NUFFT, MRI, non-Cartesian

1. INTRODUCTION

The fast evaluation of the non-uniform Fourier samples of an N-point discrete signal is a central problem in many areas including tomography [1], magnetic resonance imaging [2], synthetic aperture radar, and wavelets [3]. Since the brute force evaluation of the Fourier sum at non-uniform samples is computationally expensive, the standard approach is to obtain them as the interpolation of the uniform Fourier transform. The uniform Fourier transform is computed using the standard K-point FFT ($K \geq N$), while support limited functions (e.g., Kaiser-Bessel, Gaussian) are used to interpolate the samples. It is reported that the weighting of the signal by appropriate scale-factors, before evaluating the uniform FFT, significantly reduces the approximation error. [4, 5].

Our work is motivated by iterative non-Cartesian MRI [2]. The computational complexity and accuracy of these 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 [5]. However, 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). Although the scale-factors play a significant role in reducing the NUFFT error, they are often selected arbitrarily [4] or are restricted to parametric families with few degrees of freedom [5]. This often limits the performance of the NUFFT approximation significantly.

The main focus of this paper is to derive a memory efficient approximation to the non-uniform Fourier transform of a support-limited sequence (type 2 NUFFT). We show that the widely used NUFFT scheme is essentially a periodic shift invariant approximation of the exact discrete Fourier transform. Based on our earlier results [6], we derive an exact and computable expression for the worst-case mean square approximation error. This metric conveniently decouples the error contributions due to 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). From theoretical comparisons, we find that the OLS-NUFFT significantly improves 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 more memory efficient multi-dimensional NUFFT scheme.

2. THEORY

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 evaluate the discrete time Fourier transform (DTFT) of this sequence

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

at the non-uniform frequency locations $\nu_m; m = 0..M-1$. The direct evaluation of (1) at these samples is computationally expensive (requires $\mathcal{O}(MN)$ operations). To reduce the computational cost in evaluating (1), 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 $h[n]x[n]$:

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

The above summation is evaluated efficiently 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 [4] or as a function of φ [5]. The interpolation is performed using φ, assuming periodic boundary conditions [4, 5]. This approximation can be expressed in

two alternate ways:

$$\hat{x}_{\text{app}}(\nu) = \sum_{k\in\mathbb{Z}} c_p[k]\varphi\left(\frac{K}{N}\nu - k\right) \quad (3)$$

$$= \sum_{k=-K/2}^{K/2-1} c[k]\varphi_p\left(\frac{K}{N}\nu - k\right) \quad (4)$$

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

$$\varphi_p(\nu) = \sum_{k\in\mathbb{Z}} \varphi(\nu - kK). \quad (5)$$

In (3), $c_p[k] = \sum_{l\in\mathbb{Z}} c[k + lK]$ is the sequence obtained as the periodization of $c[k]$.

2.1. Reinterpretation of NUFFT

To better analyze the NUFFT scheme, we now reformulate the standard NUFFT approximation as the projection of the discrete Fourier transform $\hat{x}(\nu)$ in (1) on to a shift invariant space. The NUFFT scheme approximates the periodic signal $\hat{x}(\nu)$ as a linear combination as in (3), where the coefficients $c_p[k]$ are obtained as the K-point FFT of $h[n]x[n]$ as in (2):

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

$$= \int_{-\frac{N}{2}}^{\frac{N}{2}} \hat{x}(\nu) \underbrace{\sum_{n=-N/2}^{N/2-1} h[n]e^{j\frac{2\pi}{K}\left(\frac{K\nu}{N}-k\right)n}}_{\tilde{\varphi}_p^*\left(\frac{K}{N}\nu-k\right)} d\nu \quad (6)$$

In the second step, we substituted $x[n]$ by the inverse Fourier transform of $\hat{x}(\nu)$ and interchanged the order of the integration and the summation. Here $\tilde{\varphi}^*$ is the conjugate of $\tilde{\varphi}$. (6) enables us to reinterpret the derivation of the coefficients $c_p[k]$ as the evaluation of the inner-product between the discrete Fourier transform of $x[n]$ (denoted by $\hat{x}(\nu)$) and the periodized analysis functions $\tilde{\varphi}_p$. The analysis functions are specified by $\hat{\tilde{\varphi}}_p(2\pi n/K) = h[n]$. Thus, the NUFFT approximation $\hat{x}_{\text{app}}(\nu)$ is essentially a projection of the exact DFT ($\hat{x}(\nu)$) onto the shift invariant space $V_{\varphi_p} = \text{span}\left\{\varphi_p\left(\frac{K}{N}\nu - k\right); k = -K/2 \ldots K/2-1\right\}$. This re-interpretation of NUFFT enables us to use the sophisticated tools in the context of SI representations to analyze the performance of NUFFT approximation.

2.2. Quantitative expression of the error

We had derived the expression for the average error in representing an arbitrary function $s(\nu)$ in a shift invariant (SI) space in [6]. The space V_{φ_p} is only integer shift variant; shifting the function $s(\nu)$ to $s_\tau(\nu) = s(\nu - \tau)$ by non integer multiples of the sampling step affects the approximation error. The error is periodic with period $T = N/K$. The average mean square error (averaged over all possible shifts) is shown in [6] as

$$\eta(s, \varphi, \tilde{\varphi}, K) = \frac{K}{N}\int_0^{N/K} \|s_\tau(\cdot) - s_{\tau,\text{app}}(\cdot)\|^2_{L_2([-N/2,N/2))} d\tau$$

$$= \sum_{n=-\infty}^{\infty} |\hat{s}[n]|^2 E_{\varphi,\tilde{\varphi},K}\left(\frac{2\pi n}{K}\right), \quad (7)$$

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

$$E_{\varphi,\tilde{\varphi},K}(\omega) = \underbrace{1 - \frac{|\hat{\varphi}(\omega)|^2}{\hat{a}_\varphi(\omega)}}_{E_{\min}(\omega)} + \underbrace{\hat{a}_\varphi(\omega)\left|\hat{\tilde{\varphi}}(\omega) - \hat{\varphi}_d(\omega)\right|^2}_{E_{\text{res}}(\omega)}. \quad (8)$$

In (13), $\hat{s}[n]; n \in \mathbb{Z}$ are the Fourier series coefficients of $s(\nu)$, defined by

$$\hat{s}[n] = \frac{1}{N}\int_{-N/2}^{N/2} s(\nu)e^{\frac{-j2\pi\nu n}{N}} d\nu. \quad (9)$$

In (8), $\hat{a}_\varphi(\omega) = \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega + 2k\pi)|^2$ and $\hat{\varphi}_d(\omega) = \hat{\varphi}(\omega)/\hat{a}_\varphi(\omega)$ denotes the dual function of φ. Both $E_{\min}$ and E_{res} are positive terms. Moreover, $E_{\min}$ is only dependent on φ, while E_{res} is is dependent of the analysis function $\tilde{\varphi}_p$.

2.3. Error analysis of the NUFFT approximation

We now apply the error formula to analyze the NUFFT approximation. Substituting for $s(\nu)$ by $\hat{x}(\nu)$ in (13) from (1), we obtain

$$\eta(\hat{x}, \varphi, \tilde{\varphi}, K) = \sum_{n=-N/2}^{N/2-1} |x[-n]|^2 E_{\varphi,\tilde{\varphi},K}\left(\frac{2\pi n}{K}\right), \quad (10)$$

The negative sign in the index of $x[n]$ is due to the difference in the definitions of $\hat{s}[n]$ in (9) and $x[n]$ in (1). Note from (10) and (8) that the second term (due to E_{res}) will disappear if

$$\hat{\tilde{\varphi}}(2\pi n/K) = \frac{\hat{\varphi}(2\pi n/K)}{\sum_{k\in\mathbb{Z}} |\hat{\varphi}(2\pi n/K + 2k\pi)|^2}; \quad n = -N/2+1, ..N/2 \quad (11)$$

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

$$\eta_{\min}(\hat{s}, \varphi, K) = \sum_{n=-N/2+1}^{N/2} |x[n]|^2 \frac{\sum_{k\neq 0} |\hat{\varphi}(2\pi n/K + 2k\pi)|^2}{\sum_{k\in\mathbb{Z}} |\hat{\varphi}(2\pi n/K + 2k\pi)|^2} \quad (12)$$

Here we used the symmetry of the error kernel. The above expression indicates 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 (12) is independent on the analysis function $\tilde{\varphi}$. The approximation $s_{\text{app}}(\nu)$, derived using of the optimal analysis function specified by (11), is the orthogonal projection of $s(\nu)$ onto V_{φ_p}.

The use of the optimal scale factors (11) will minimize the approximation error for a specified interpolating kernel. More importantly, this choice will lead to an error expression that is independent of the scale factors. By evaluating the worst case least square error, we will now obtain an error expression that is also independent of the signal $x[n]$.

$$\eta^2_{\max} = \arg\max_{x:\|x\|_{L_4}=1} \eta^2_{\min} = \sum_{n=-N/2}^{N/2-1} \left[E_{\min}\left(\frac{2\pi n}{K}\right)\right]^2$$

$$= \sum_{n=-N/2}^{N/2-1} \frac{\left(\sum_{k\neq 0} \left|\varphi(\frac{2\pi n}{K} + 2k\pi)\right|^2\right)^2}{\left(\sum_{k\in\mathbb{Z}} \left|\varphi(\frac{2\pi n}{K} + 2k\pi)\right|^2\right)^2} \quad (13)$$

Here, we used the Cauchy Shwartz inequality. Note that this is a tight bound; any sequence that satisfies $|x[n]|^2 = E_{\min}\left(\frac{2\pi n}{K}\right)$ will lead to the worst case LS error.

2.4. Optimized least-squares (OLS) interpolator

Our goal is to derive a $\varphi : \|\varphi\|_{L_2} = 1$, which is finitely supported in the range $[-J/2, J/2]$, (i.e., $\varphi \in L_2([-J/2, J/2])$) that minimize the worst case LS error:

$$\varphi_{\text{OLS}} = \arg \min_{\varphi \in L_2[-J/2, J/2]} \sum_{n=-N/2+1}^{N/2} \frac{\left(\sum_{k\neq 0} |\varphi(x+2k\pi)|^2\right)^2}{\left(\sum_{k\in\mathbb{Z}} |\varphi(x+2k\pi)|^2\right)^2}. \tag{14}$$

Since this metric is a non-linear functional of φ, it is difficult, if not impossible, to derive the optimal continuous domain function φ_{OLS}. We propose derive the optimal discretization of φ using an iterative re-weighted quadratic minimization. Towards this end, we rewrite the the error expression as

$$\begin{aligned}\eta^2 &= \sum_{n=-N/2}^{N/2-1} \underbrace{\frac{\sum_{k\neq 0} \left|\hat{\varphi}\left(\frac{2\pi(n+kK)}{K}\right)\right|^2}{\left(\sum_k \left|\hat{\varphi}\left(\frac{2\pi(n+kK)}{K}\right)\right|^2\right)^2}}_{h_\varphi[n]} \sum_{k\neq 0} \left|\hat{\varphi}\left(\frac{2\pi(n+kK)}{K}\right)\right|^2 \\ &= \sum_{n\in\mathbb{Z}} g_\varphi[n] \left|\hat{\varphi}\left(\frac{2\pi n}{K}\right)\right|^2 \end{aligned} \tag{15}$$

where

$$g_\varphi[n] = \begin{cases} 0; & -K/2 < n < K/2 - 1 \\ h_\varphi[n - 2lK]; & (lK - N/2) < n < (lK + N/2 - 1) \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Assuming the weights $g[n]; n \in \mathbb{Z}$ to be specified, the criterion (15) is a simple weighted norm in the Fourier domain. We will use this simplification to derive the optimal interpolator in the next subsection.

2.5. Algorithm

We propose a two step iterative algorithm that uses the following steps: Start with an initial set of weights $g[n]$, corresponding to $h[n] = 1; n = -N/2..N/2 - 1$.

1. Using the current value of $g[n]$, derive the optimal φ that minimizes (15), subject to $\|\varphi\|_{L_2} = 1$.
2. Using the current value of φ, evaluate the optimal weights $g_\varphi[n]$ specified by (16).
3. Exit if converged. Else, goto step 1.

We will now focus on step (1), where we derive the optimal discretized φ that minimize (15) subject to $\|\varphi\|^2$=1. Assuming an oversampling factor O and by choosing $q[n] = \varphi(n/O)$, we reformulate the derivation as

$$\mathbf{q}_{\text{OLS}} = \arg\min_{\mathbf{q}} \left(\mathbf{q}^T \underbrace{\mathbf{F}^T \mathbf{W}_g \mathbf{F}}_{\mathbf{A}} \mathbf{q} \right) \text{ subject to } \mathbf{q}^T\mathbf{q} = 1, \tag{17}$$

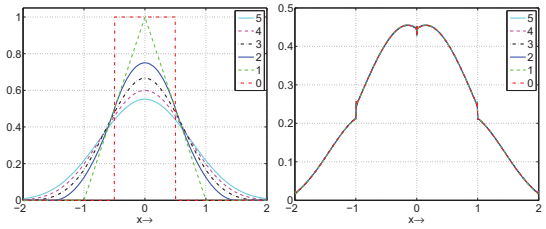
where $\mathbf{F}$ is the $KO \times J/T - 1$ DFT matrix, $\mathbf{W}_g$ is the $KO \times KO$ matrix with diagonal entries corresponding to $g_\varphi[n]$. Note that $\mathbf{A}$ is a small matrix of dimension $OJ - 1 \times OJ - 1$. We solve this constrained minimization problem using Lagrange's multiplier methods to obtain

$$\mathbf{A}\mathbf{q}_{\min} = \lambda_{\min}\mathbf{q}_{\min} \tag{18}$$

where $\lambda_{\min}$ is the minimum eigen value of $\mathbf{A}$ and $\mathbf{q}_{\min}$ is the corresponding eigen vector. Since the iterative two-step process is an approximation to the minimization of the non-linear criterion η^2, we use a relaxation to stabilize the algorithm. Specifically, we specify the optimal $\mathbf{q}$ at the iteration i, denoted by $\mathbf{q}_i$ as $\mathbf{q}_i = \alpha_i\, \mathbf{q}_{\min} + (1 - \alpha_i)\, \mathbf{q}_{i-1}; 0 \le \alpha_i \le 1$,where $\mathbf{q}_{\min}$ is the solution to (18) at that step and $\alpha_i = \arg\min_\alpha \left(\eta^2 \left(\alpha\, \mathbf{q}_{\min} + (1 - \alpha)\, \mathbf{q}_{i-1}\right)\right)$. This approach is analogous to a line search; it ensures that the iterative algorithm monotonically decrease the worst-case errors. At each iteration, we check for the value of α_i and exit if it is below a specified threshold.

3. RESULTS

We first demonstrate the convergence of the algorithm in a specific example ($K = 132, N = 128, O = 100, J = 4$). We consider a wide range of input initializations (ranging from Bspline of order 0 to order 5). It is seen from Fig. 1 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.



(a) initializations (b) final solution

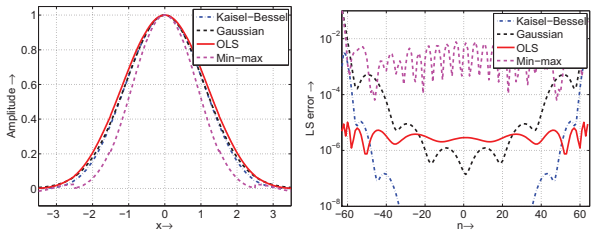
Fig. 1. Illustration of the convergence of the algorithm. The worst-case errors are also the same, up to numerical precision.

3.1. Comparison with standard interpolators

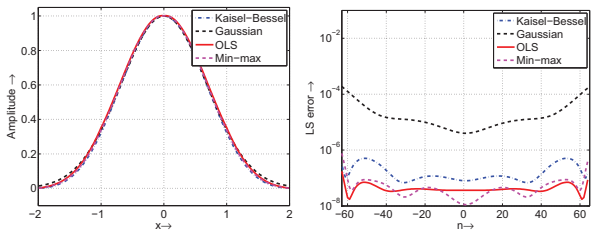
We now compare the performance of the OLS function with Kaiser-Bessel, Gaussian and min-max interpolators in this subsection. To make the comparisons fair, we optimized these interpolators with respect to the mean square criterion. We also use the OLS scale-factors (11) for these interpolators. For the min-max interpolators with $K = 2N$, we employ the scale-factors with the highest order that are reported in [5]. Since the optimal min-max scale-factors were only derived in [5] for $K = 2N$, we obtained the scale factors by performing least square fit to the optimal Kaiser-Bessel scale-factors (with $\beta = 1; L = 22$) as in [5].

The comparison of the interpolators and the error kernels at $K = 132$ and $K = 256$ are shown in Fig. 2. It is seen that the subtle variations in the shape of the interpolators in Fig. 2-a lead to significant differences 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. 2-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).

In Fig. 3, we compare the different interpolator families based on the worst-case mean square error (13). The comparisons were performed at $K = 132$, $K = 140$ and $K = 256$ respectively. Fig. 3.a denote the error curves at K=132 and oversampling factor=100. Similar comparisons are shown for $K = 140$ and $K = 256$ in Fig. 3.b and Fig. 3.c respectively. The min-max estimator provides

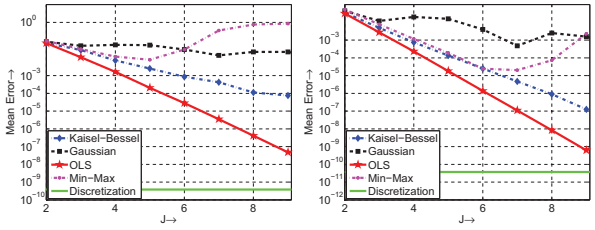


(a) interpolator: $K = 132; J = 7$ (b) MS kernel: $K = 132; J = 7$

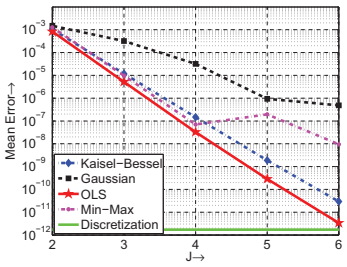


(c) interpolator: $K = 256; J = 4$ (d) MS kernel: $K = 256; J = 4$

Fig. 2. Comparison of the OLS interpolator with classical schemes. (a)-(c) compares the interpolators at $K = 132$, $N = 128$,$O = 100$ and $J = 7$. (a) indicates the shape of the interpolators (b). Comparison based on the mean square error kernel. (c). Comparison based on the min-max error criterion derived in [5]. (d)-(f). Comparison of the interpolators for $K = 256$, $N = 128$, $J = 4$ 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 respectively.



(a) K=132 (b) K=140



(c) K=256

Fig. 3. Theoretical comparison of the different interpolator families. (a)&(d) indicate the decay of the worst case mean square and min-max errors respectively for K=132 and oversampling factor=100. The worst-case discretization error is displayed in magenta. (b)&(e) 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 [5] were used for $K = 2N$. (c) Comparison of different kernels at K=256, N=128 and $O = 190$.

comparable errors for lower values of J. However as the length of the interpolator increase, its performance deteriorate. This is due to the limited number of scale-factor parameters that can be derived in the min-max setting.

From these comparisons, it is seen that the proposed OLS NUFFT scheme significantly outperforms its closest competitor: the Kaiser Bessel interpolator, especially when K is small. For example, the use of the OLS interpolator provides approximately a factor of 5×10^3 decrease in the mean-square error at $J = 9; K = 132$. The interpolator settings (J=9;K=132) provides a worst-case error of $3{*}10^{-8}$ that 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.

4. CONCLUSION

In this paper, we designed a new interpolator and its optimal scale factors for the efficient evaluation of non-uniform Fourier transform. Using the expression for the average least squares error in the context of periodic shift invariant representations, we derived the optimal scale factors for a specified interpolator. Assuming the use of these scale factors, we derived a worst-case error criterion that is only dependent on the interpolating function. We then derived the optimized least square NUFFT interpolator using an iterative reweighted optimization algorithm. The significant performance improvement provided by this scheme for low oversampling factors could lead to faster and more memory efficient conjugate algorithms for non-Cartesian Fourier inversion.

5. REFERENCES

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

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

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

[4] 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.

[5] J. Fessler and B. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Transactions on Signal Processing*, Jan 2003.

[6] M. Jacob, T. Blu, and M. Unser, "Sampling of periodic signals: a quantitative error analysis," *IEEE Transactions on Signal Processing*, Jan 2002.