# On the Optimality of the Gridding Reconstruction Algorithm

Hossein Sedarat* and Dwight G. Nishimura

***Abstract*—Gridding reconstruction is a method to reconstruct data onto a Cartesian grid from a set of nonuniformly sampled measurements. This method is appreciated for being robust and computationally fast. However, it lacks solid analysis and design tools to quantify or minimize the reconstruction error. Least squares reconstruction (LSR), on the other hand, is another method which is optimal in the sense that it minimizes the reconstruction error. This method is computationally intensive and, in many cases, sensitive to measurement noise. Hence, it is rarely used in practice.**

**Despite their seemingly different approaches, the gridding and LSR methods are shown to be closely related. The similarity between these two methods is accentuated when they are properly expressed in a common matrix form. It is shown that the gridding algorithm can be considered an approximation to the least squares method. The optimal gridding parameters are defined as the ones which yield the minimum approximation error. These parameters are calculated by minimizing the norm of an approximation error matrix. This problem is studied and solved in the general form of approximation using linearly structured matrices.**

**This method not only supports more general forms of the gridding algorithm, it can also be used to accelerate the reconstruction techniques from incomplete data. The application of this method to a case of two-dimensional (2-D) spiral magnetic resonance imaging shows a reduction of more than 4 dB in the average reconstruction error.**

***Index Terms*—Gridding reconstruction, image reconstruction, matrix approximation, nonuniform sampling.**

## I. Introduction

THE problem of reconstructing data onto a Cartesian grid from a set of nonuniformly sampled measurements arises in many disciplines ranging from radio astronomy [1] to various modalities of medical imaging [2]–[4]. In addition to some algorithms developed for specific sampling patterns, such as filtered back projection for radial sampling [5], [6], many other methods try to solve this problem for a general sampling pattern. These methods are based on some form of interpolation, such as nearest neighbor, bilinear transform, local averaging, and fixed kernel interpolation [7]–[11]. Perhaps the most general method in this category is the gridding reconstruction algorithm [12], [13]. This method is computationally fast, robust to measurement errors, and intuitively appealing. Hence, it is widely used in practice.

Manuscript received September 28, 1999; revised January 30, 2000. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was Z. Liang. *Asterisk indicates corresponding author*.

*H. Sedarat and D. G. Nishimura are with the Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA 94305-9510 USA (e-mail: sedarat@lad.stanford.edu).

Publisher Item Identifier S 0278-0062(00)04503-1.

The gridding algorithm consists of two main steps: normalization and local averaging. In the first step, the measured data is normalized to compensate for the nonuniform sampling density. The normalization factors, known as density compensation factors (DCF's), are inversely proportional to the local sampling density. Therefore, the measurements taken from densely sampled regions are multiplied by small numbers and those from sparsely sampled regions are multiplied by large numbers. This normalization scheme ensures that despite the nonuniform sampling pattern, the measurement space is uniformly weighted. In the second step, the grid data is calculated by averaging the neighboring samples. This step is formally expressed as the convolution of the normalized data with an interpolating kernel. The interpolating kernel is normally chosen to have a bell shape to emphasize the contribution of the samples closer to the grid points.

Much of the research in the area of gridding reconstruction has concentrated on the calculation of DCF's based on various definitions of the local sampling density. Some of the proposed methods derive the sampling density from the analytical expression of the trajectory along which the samples are taken [14], [15]. Later work in this area allows multiple trajectories with crossovers [16], [17]. In another approach, the space of measurements is transformed to another space in which the sampling pattern is uniform [18]. The determinant of the Jacobian of the transform operator is used as a measure of the local sampling density. There are also other methods in which the sampling space is partitioned into several cells, each cell representing the neighborhood associated with a sample point. The area of each cell is used as the DCF for the corresponding sample point [19]–[21]. More recently and based on the work described in [13], Pipe *et al.* [22] have suggested an iterative algorithm which tries to satisfy a necessary condition for correct reconstruction.

An efficient implementation of the gridding algorithm requires an interpolating kernel with narrow width. Interpolation with any kernel other than a $sinc(\cdot)$ function results in a magnitude modulation in the Fourier domain. This effect is compensated by multiplying the Fourier transform of the data by a proper deapodization function. Jackson *et al.* [13] have extensively studied various kernels and compared them in terms of the relative amount of aliasing energy that they deposit inside the field of view (FOV). A more recent study suggests that a better performance may be achieved by fine tuning the kernel according to the support region of the data [23].

None of these studies have rigorously addressed the effect of various gridding parameters on the reconstruction error. Neither is it clear how to adjust these parameters to achieve the minimum reconstruction error. In contrast to the gridding algo-

rithm, there are other reconstruction techniques that are explicitly designed to minimize the energy of the reconstruction error [24], [25]. These so-called least squares methods take advantage of the linear relationship between the grid data and the sampled points to find the minimum error solution. Least squares techniques are normally computationally complex and typically ill-conditioned for highly dense sampling patterns and, thus, are sensitive to noise and measurement errors.

In this paper, we establish the relationship between the gridding algorithm and the least squares method. These two reconstruction techniques reveal their similarities when they are expressed in a common matrix form. We show that the gridding algorithm can be derived from the least squares method by proper matrix approximation. This approximation is intended to simplify the implementation of the gridding algorithm by replacing the unstructured matrices in the least squares method with diagonal or band matrices. The optimal gridding parameters are obtained by minimizing the average error of approximation. We present a method to solve this problem in the more general form of approximation using linearly structured matrices.

This method is applied directly to derive a closed-form solution for the optimal DCF's. This derivation requires only information about the sampling coordinates and does not rely on the analytical expression of the sampling trajectory. An interpretation of the optimal compensation factors indicates that these factors are indeed inversely related to the sampling density and suggests a new approach in the calculation of the local sampling density. We show that our proposed method can be easily extended to minimize a weighted average of the approximation error. The weighting function can be chosen to emphasize a subspace of the acquisition space or for a better reconstruction of a portion of the grid data.

We also present a method to obtain the optimal interpolating kernel and deapodization function. The joint optimization of these two functions is inherently a nonlinear problem and rather hard to solve directly. We propose an iterative algorithm that reduces this problem to a linear problem in each step allowing the application of our matrix approximation technique. Our methodology can be used to generalize the gridding algorithm. We study examples of such cases in interpolation with a shift-variant kernel and extensions of gridding in reconstruction from incomplete data. Finally, we present some of our results from a case of 2-D spiral magnetic resonance (MR) imaging.

## II. Background

This section presents an overview of both the gridding and the least squares reconstruction (LSR) methods. The goal is to express these two techniques in a common frame work and thereby establish their relationship. Although the discussion is presented in the context of data acquisition and reconstruction in MR imaging, the final results hold for other applications as well.

MR raw data consists of several measurements of the image in the Fourier domain, or better known in the MR community as $k$-space [4], [26]. These measurements are discrete samples of the Fourier transform of the image taken over a sampling trajectory in $k$-space. Fig. 1(a)–(c) shows three trajectories widely

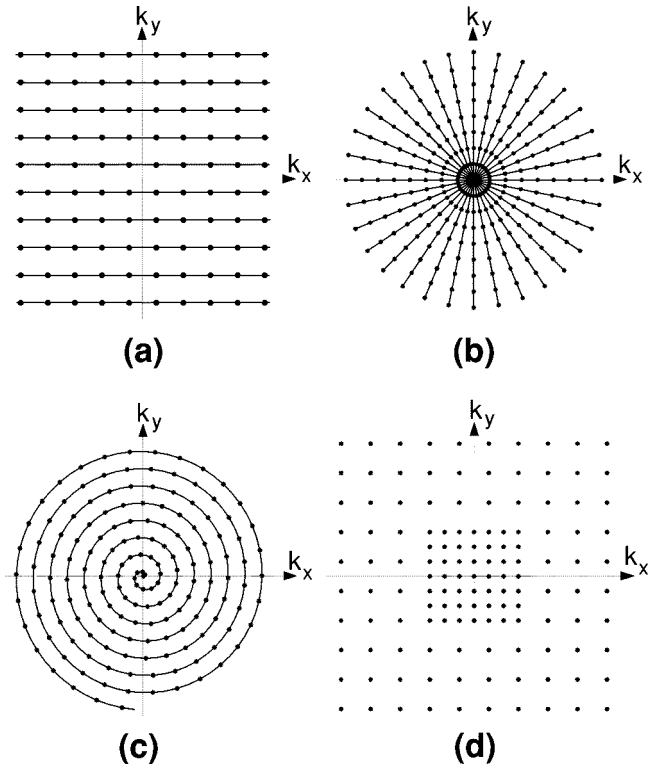


Fig. 1. The sampling patterns known as (a) 2DFT, (b) PR, and (c) spiral are commonly used in 2D MR imaging. (d) Variable-density 2DFT (VD-2DFT) is another nonuniform sampling pattern with predictable properties but no practical applications.

used in 2-D MR imaging: 2-D Fourier transform (2DFT), projection reconstruction (PR), and spiral. With a uniform sampling pattern, as in the case of the 2DFT trajectory, the image can be reconstructed easily by taking the inverse FFT of the raw data. However, when $k$-space is not uniformly sampled, as in PR and spiral, the reconstruction algorithm needs to derive the grid points from the acquired data prior to the inverse FFT operation.

Fig. 1(d) shows another nonuniform sampling pattern which is very similar to a 2DFT trajectory, except for the central section of $k$-space where the sampling density is higher. We refer to this sampling pattern as the variable-density 2DFT trajectory (VD-2DFT). Although this sampling pattern has never been used in practice, it provides a simple example with predictable properties that is used to verify some of the results of this paper.

With a finite imaging FOV, the Nyquist sampling theorem relates the acquired trajectory data to the desired grid points as

$$m(\kappa_i) = \sum_{j=1}^{N} \mathrm{sinc}(\kappa_i - k_j) m(k_j) \tag{1}$$

where $m(\cdot)$ is the data in $k$-space, $\{\kappa_i, i = 1, \ldots, L\}$ is the set of trajectory points, and $\{k_j, j = 1, \ldots, N\}$ is the set of grid points. For simplicity, all $k$-space points are normalized to the FOV. It is assumed that the image has limited resolution, thus a finite number of grid points at the center of $k$-space can represent the image uniquely. This assumption may theoretically contradict the earlier assumption of having an image with finite FOV. However, with a proper coverage of $k$-space, the grid

points at higher spatial frequencies can justifiably be ignored in many practical cases where the energy of the image is dominantly concentrated at the center of $k$-space [27]. Normally, the number of trajectory points $(L)$ is not smaller than the size of the grid $(N)$.

Equation (1) provides a method to interpolate the trajectory data from the grid samples. The problem of image reconstruction in MRI is the inverse: the available data is the nonuniform measured data over the $k$-space trajectory $(m(\kappa_i))$ and the unknowns are the samples on the rectangular grid $(m(k_j))$. The gridding reconstruction method tries to solve this problem by interpolating the grid points back from the trajectory data. In its simplest form, this method uses the same interpolating kernel used in (1), i.e., the $\mathrm{sinc}(\cdot)$ function. However, to take into account the nonuniform sampling density of $k$-space, the measurements are normalized by the so-called DCF's prior to the interpolation.

The gridding operation can be expressed as

$$\hat{m}(k_j) = \sum_{i=1}^{L} \mathrm{sinc}(k_j - \kappa_i) m(\kappa_i) d(\kappa_i) \tag{2}$$

where $\hat{m}(k_j)$ denotes the calculated grid data and $d(\kappa_i)$ is the DCF at trajectory point $\kappa_i$. The DCF's are inversely proportional to the local sampling density at their corresponding $k$-space points.

Another method of reconstruction is least squares reconstruction (LSR). Equation (1) indicates a linear relationship between the measured trajectory data and the unknown grid points. This linear relationship can be easily inverted to obtain the grid data from the measured samples. Let $\boldsymbol{m} = [m_i]$ denote the $N \times 1$ vector of grid points with $m_i = m(k_i)$ and $\boldsymbol{\mu} = [\mu_i]$ denote the $L \times 1$ measurement vector with $\mu_i = m(\kappa_i)$. We can rewrite (1) as

$$\boldsymbol{\mu} = T\boldsymbol{m} \tag{3}$$

where $T = [t_{ij}]$ is the $L \times N$ interpolation matrix with $t_{ij} = \mathrm{sinc}(\kappa_i - k_j)$. Note that the $j$th column of $T$ consists of the trajectory samples of the interpolating kernel centered at the $j$th grid point $k_j$.

LSR solves this linear equation for a grid vector that yields the minimum measurement error. This solution can be expressed in terms of the pseudoinverse of $T$ derived from the singular values of that matrix [28]. When the interpolation matrix is full rank this solution can be expressed as

$$\hat{\boldsymbol{m}} = (T^*T)^{-1}T^*\boldsymbol{\mu} \tag{4}$$

where $T^*$ denotes the complex conjugate transpose of the matrix $T$. This method of reconstruction is optimal in the sense that it results in minimum error, hence the name least-squares reconstruction. LSR can be ill-conditioned with highly over-sampled trajectories. In such cases, the problem should be regularized by taking into account the effect of noise and measurement error.

The LSR and the gridding reconstruction equations (2) and (4) are obtained by different approaches and they seemingly have different forms. Nevertheless, these two reconstruction techniques are closely related. To establish this relationship, the gridding reconstruction of (2) is rewritten in a matrix form as

$$\hat{\boldsymbol{m}} = T^*D\boldsymbol{\mu} \tag{5}$$

where $\hat{\boldsymbol{m}}$, $\boldsymbol{\mu}$ and $T$ are defined as before and $D_{L\times L} = [d_{ij}]$ is the diagonal density compensation matrix with the DCF's as its diagonal elements: $d_{ii} = d(\kappa_i)$.

To obtain a similar form for LSR (4) is rewritten as

$$\hat{\boldsymbol{m}} = T^*P\boldsymbol{\mu} \tag{6}$$

where the compensation matrix $P_{L\times L}$ is defined as $P = T(T^*T)^{-2}T^*$. Comparing (5) and (6), it is clear that gridding and LSR have very similar forms. In fact, one can consider the gridding algorithm as a special form of LSR where the compensation matrix $P$ is approximated with a diagonal matrix $D$. This approximation allows fast calculation of the compensated data. This calculation has a complexity on the order of $L^2$ in LSR and only on the order of $L$ in gridding ($L$ is the number of trajectory samples).

## III. Optimal Density Compensation

Up to this point, the relationship between the gridding reconstruction algorithm and the LSR method has been established. Moreover, it has been shown that the gridding method can be considered as an approximation to LSR. To introduce the notion of optimality to gridding reconstruction, one needs to show how to do this approximation optimally. In this section, we propose a method to obtain the optimal DCF's such that the error due to the gridding approximation is minimum.

The reconstruction error due to the gridding approximation is

$$\boldsymbol{e} = T^*(P - D)\boldsymbol{\mu}. \tag{7}$$

The goal is to find a diagonal matrix $D$ that optimally approximates the compensation matrix $P$ in the sense that the norm of error vector $\mathbf{e}$ is minimized. The error vector depends not only on the choice of matrix $D$ but also on the vector of measurements $\mu$. To minimize the error for all possible measurements, the norm of error matrix $E = T^*(P-D)$ has to be minimized. Note that, by definition, the norm of a matrix indicates how much the length of a vector is amplified when multiplied by that matrix [28]. Thus, by minimizing the norm of the error matrix $T^*(P-D)$, one can assure that the approximation error will be minimum for all possible measurement vectors.

Appendix I presents a method on optimal matrix approximation. We apply the method described there to find the optimal density compensation matrix $D$. The final result is

$$d_{ii} = \frac{[T(T^*T)^{-1}T^*]_{ii}}{[TT^*]_{ii}}, \qquad i = 1 \ldots L \tag{8}$$

and the average squared error of approximation is obtained as

$$\epsilon = \sum_{i=1}^{L} [T(T^*T)^{-2}T^*]_{ii} - \sum_{i=1}^{L} d_{ii}^2 [TT^*]_{ii} \tag{9}$$

where $[A]_{ij}$ and $a_{ij}$ both denote the $(i,j)^{\mathrm{th}}$ element of matrix $A$.

To ensure that the solution of (8) always exists, we need to show that the denominator is nonzero and the numerator is well defined. The $i^{\text{th}}$ diagonal elements of $TT^*$ can be expressed as

$$[TT^*]_{ii} = \sum_{j=1}^{N} \text{sinc}^2(\kappa_i - k_j). \tag{10}$$

This is the amount of energy that a $\text{sinc}(\cdot)$ kernel centered at the $i^{\text{th}}$ trajectory point deposits on the grid points. This factor is always greater than zero and, more specifically, for a grid of infinite extent this factor is equal to unity. For a finite grid size and for sample points close to the center of $k$-space these values are very close to one, but for samples at the edge of $k$-space they may be smaller than one but always greater than zero.

The other term of concern is $(T^*T)^{-1}$ in the numerator of (8). When the condition number of the interpolation matrix $T$ is large, which may be the case for highly over sampled trajectories, the calculation of $(T^*T)^{-1}$ is numerically unstable. However, the whole numerator, which represents a projection operator to the range space of the matrix $T$, is independent of the condition number of that matrix. To confirm this fact, we start with the singular value decomposition of matrix $T$ as

$$T = \Theta \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \Omega^*$$

where $\Theta_{L\times L}$ and $\Omega_{N\times N}$ are unitary matrices and $\Sigma$ is a diagonal matrix with the singular values of $T$ on the diagonal. After some basic matrix manipulations, the numerator of (8) can be re-expressed as

$$T(T^*T)^{-1}T^* = \Theta \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \Theta^*$$

where $I$ is an identity matrix with dimensions similar to $\Sigma$. This equation indicates that the numerator in (8) is independent of the singular values of $T$. It also presents a stable method to calculate that numerator.

The optimal DCF's are calculated for the VD-2DFT, PR, and spiral trajectories of Fig. 1. We assume an oversampling factor of four at the center of $k$-space in the VD-2DFT trajectory. The optimal DCF's for this trajectory is shown in Fig. 2(a). These factors are close to unity everywhere except at the center of $k$-space where they take values close to $1/4$. As expected, these values are inversely proportional to the local sampling density.

For a PR trajectory with fine sampling pattern, the local sampling density is inversely proportional to the distance from the $k$-space origin. Therefore, the DCF's at each $k$-space point are conventionally chosen to be proportional to the distance of that point from the origin. This fact is also well known from the theory of tomography and the Radon transform [29]. The three-dimensional 3-D contour plot of Fig. 2(b) shows that the optimal DCF's also endorse this linear relationship. The small variations of the optimal DCF's for points on a constant radius is due to the fact that unlike the sampling pattern of PR trajectory, the Cartesian grid pattern does not have a circular symmetry. Consequently, the average distance between trajectory samples and grid points is angle dependent. Hence, the optimal DCF's can vary for points on a constant radius. This variation is smaller at the center of $k$-space where the sampling density is higher and larger at the periphery of $k$-space where the local sampling density is lower.

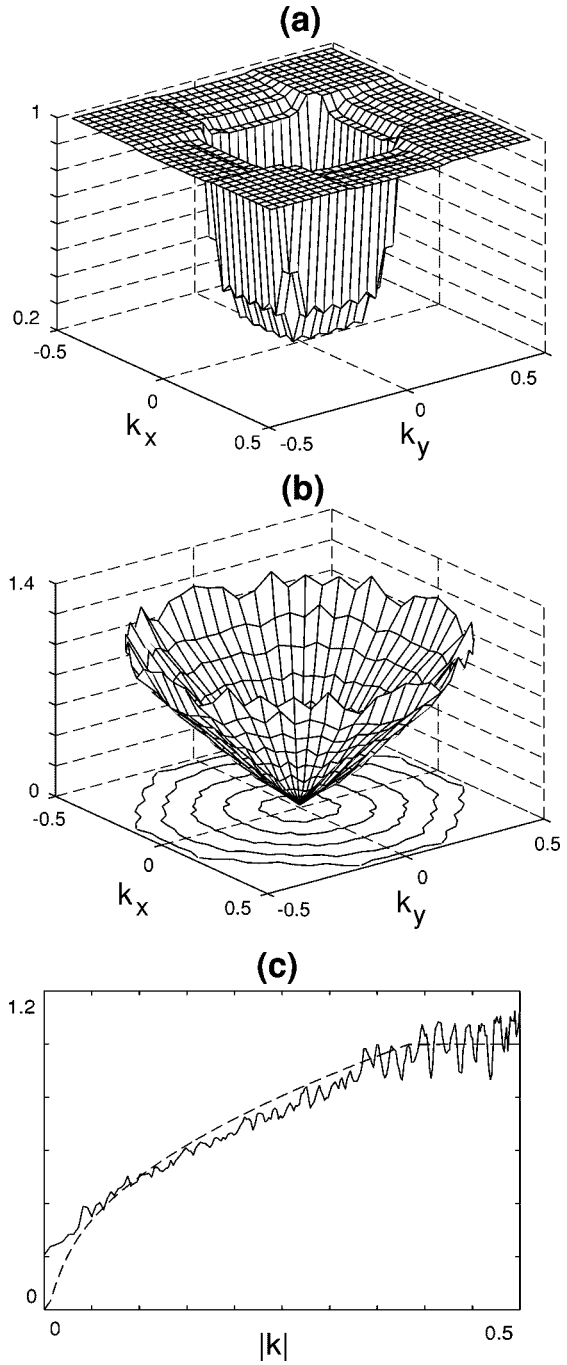


Fig. 2. The optimal DCF's for (a) VD-2DFT and (b) PR trajectories (contour plot). (c) The optimal (solid) and conventional (dashed) DCF's for the spiral trajectory.

The optimal DCF's for a spiral trajectory are plotted in Fig. 2(c). This figure shows that the conventional DCF's obtained analytically by Meyer *et al.* [16] are very close to the optimal values. Note the difference between the two DCF's at the $k$-space origin. While the analytical expression assumes an infinite sampling density at origin resulting in a zero DCF, the optimal DCF is nonzero complying with the fact that the sampling density at the origin is finite.

### A. Interpretation

In the previous section, we presented a method to obtain the DCF's that are optimal in the sense that they minimize the average approximation error in the gridding reconstruction algorithm. We also showed that, for some typical trajectories, these factors are very close to what is expected and also to what is obtained through other conventional methods. However, we have not shown that the optimal factors obtained from (8) carry the notion of density compensation in any way. This section presents an argument to indicate that the optimal DCF's are in fact inversely related to the local sampling density. The conventional methods for calculating the DCF's use various definitions for sampling density, all of which are intuitively appealing. The argument in this section leads to yet another definition for local sampling density which is derived from the optimization procedure of the previous section.

We start with (8). As discussed in the previous section, the denominator in this equation is equal or very close to one. Therefore, for simplicity, we discard this term and rewrite this equation as

$$\begin{aligned} d_{ii} &\approx [T(T^*T)^{-1}T^*]_{ii} \\ &= [T(T^*T)^{-1}T^*\mathbf{u_i}]_i \end{aligned}$$

where $\boldsymbol{u_i}$ is the unity vector with all zero elements except for a 1 at the $i^{\text{th}}$ position. The matrix $T(T^*T)^{-1}T^*$ represents a projection operator which projects the vector $\boldsymbol{u_i}$ into the range space of the matrix $T$. This equation indicates that the $i^{\text{th}}$ DCF can be obtained by projecting the corresponding unity vector into the range space of $T$ and taking the $i^{\text{th}}$ element of the projected vector. The range space of $T$ is nothing but all the possible measurements that satisfy the FOV constraint in the object domain. In $k$-space, this range space consists of all signals that have smooth variations such that their inverse Fourier transform have a finite support confined in the FOV. Therefore, the operation of projection into the range of $T$ is equivalent to a low-pass filtering in $k$-space. On the other hand, the unity vector $\boldsymbol{u_i}$ represents a Kronecker delta in $k$-space centered at $\kappa_i$. Thus, the optimal DCF at a certain point in $k$-space is obtained by a proper low-pass filtering of a delta function centered at that point and measuring the amplitude of the resulting waveform at the same point.

To illustrate these operations, we use a simple one-dimensional (1-D) trajectory with the sampling pattern shown in Fig. 3(a). The local sampling density for this trajectory is high at the center and approaches unity at the periphery of $k$-space. The $k$-space velocity for this trajectory is shown in Fig. 3(b) and is assumed to have a saturated exponential profile. To obtain the DCF at the origin, we start with the delta function of Fig. 4(a) and project that to the range of $T$. The resulting waveform is the low-pass signal shown in Fig. 4(b). The corresponding DCF is the amplitude of this signal at the origin denoted as $d_0$. Unlike the original delta function, the low-pass waveform cannot have sharp variations. Thus, the final waveform shows smooth transition from large samples close to the $k$-space origin to small samples farther from origin.

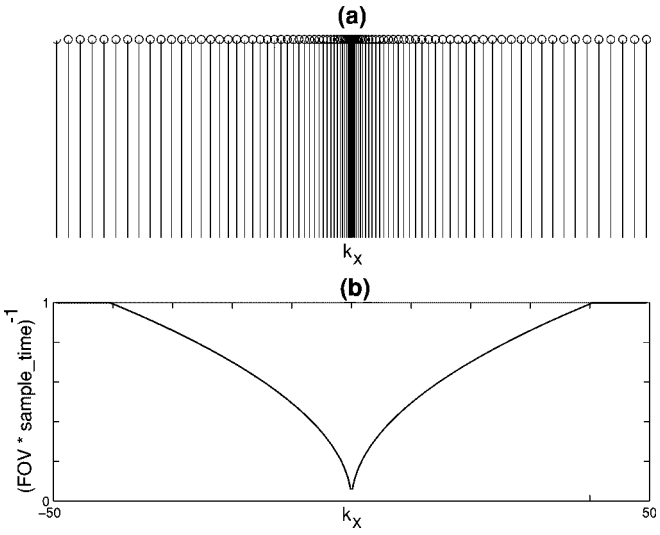


Fig. 3. (a) A nonuniform 1-D sampling pattern. (b) The corresponding $k$-space velocity.

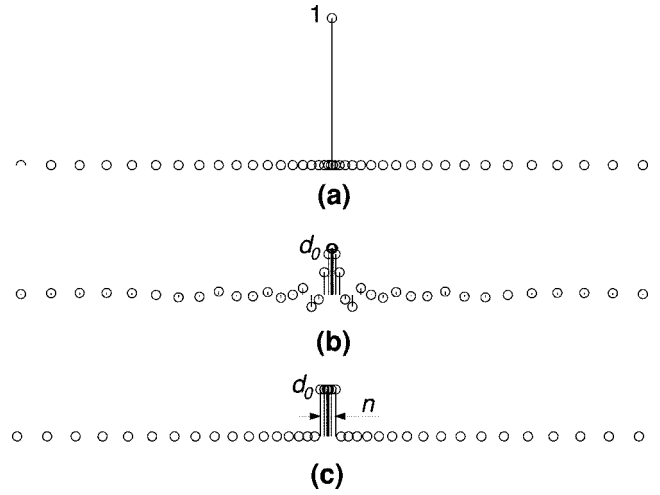


Fig. 4. To calculate the DCF at a point in $k$-space, a delta function center at that point (a) is projected into the space of $k$-space signals which satisfy the FOV constraint. The projected function (b) has a low-pass waveform which can be approximated with a rectangular function with $n$ nonzero samples (c). The amplitude of this function $(d_0)$ represent the DCF.

The number of large samples is greater when the local sampling density at origin is higher. We approximate this low-pass signal with a rectangular waveform shown in Fig. 4(c). This waveform exhibits a constant amplitude for $n$ samples close to the origin and zero for other samples. The number of nonzero samples $n$ is directly related to the sampling density at origin. The projected signal not only satisfy the FOV constraint, but it also is the closest signal, in Euclidean distance, to the original delta function. This distance for the approximated waveform is easily expressed as $(1-d_0)^2+(n-1)d_0^2$. This distance is minimized for $d_0 = 1/n$ indicating that the optimal DCF's are in fact inversely related to the local sampling density.

This discussion not only justifies the fact that the conventional DCF's are close to the optimal values, but it also provides a new definition of local sampling density which yields the optimal DCF's.

### B. Generalized Compensation Factors

In the previous section, we proposed a method to calculate DCF's such that the average approximation error in the gridding algorithm is minimized. This method is easily generalizable and can be used to minimize a weighted average of the approximation error. This is done by minimizing the norm of a weighted error matrix, $E_w$, which is obtained by pre- or post-multiplying the error matrix of (7) by some weighting matrices as

$$E_w = W_l T^*(P - D)W_r.$$

$W_l$ is the left weighting matrix that provides weighting for the grid data and can be used, for instance, to emphasize a particular region of interest in the image and reduce the reconstruction error in that region. The right weighting matrix $W_r$ introduces weighting into the measured data and can be used, for instance, to impose some constraints on the measured data. The optimal diagonal matrix $D$, which is obtained by minimizing the norm of weighted error matrix, is called the weighted compensation matrix and its diagonal elements are referred to as weighted compensation factors (WCF).

A weighting scheme of particular interest is with $W_l = I$ and $W_r = T$. With these weighting matrices the image is uniformly weighted in the object domain but the measured data is weighted with the interpolation matrix $T$. As discussed before, this matrix spans the space of all possible measurement vectors that satisfy the FOV constraint in the object domain. Therefore, with this weighting scheme the error is minimized for the measurement vectors that can be obtained from objects with the specified FOV while ignoring all other vectors. This weighting scheme also minimizes the average reconstruction error in the point spread function (PSF). This is because each column of $T^*PT$ represents the PSF at the corresponding grid point and thus the weighted error matrix $T^*(P-D)T$ represents the gridding error in the PSF. It is worth noting that the WCF's obtained from this weighting scheme are the ones discussed by Rosenfeld [25].

Using the results of Appendix I, we can show that the WCF's for this particular weighting scheme are the solution to the following linear equation:

$$A\boldsymbol{d} = \boldsymbol{b}$$

where $a_{ij} = |[TT^*]_{ij}|^2$ and $b_i = [TT^*]_{ii}$. The solution to this problem can be ill-conditioned, in which case some form of regularization may be necessary for a numerically stable solution.

These compensation factors have been calculated for the VD-2DFT trajectory. Note that, the set of sample points in VD-2DFT includes all the grid points. The calculated WCF for the trajectory points that coincide with grid points is one and the WCF for all the other points is zero. This result is expected because, in the absence of noise, knowing the sampling points on the grid can uniquely express the image and the extra trajectory points off the grid do not add any extra information. Therefore, with these compensation factors, the image is reconstructed perfectly and the error in the PSF will be zero. This also explains the fact that the approximation error using these WCF's is zero.

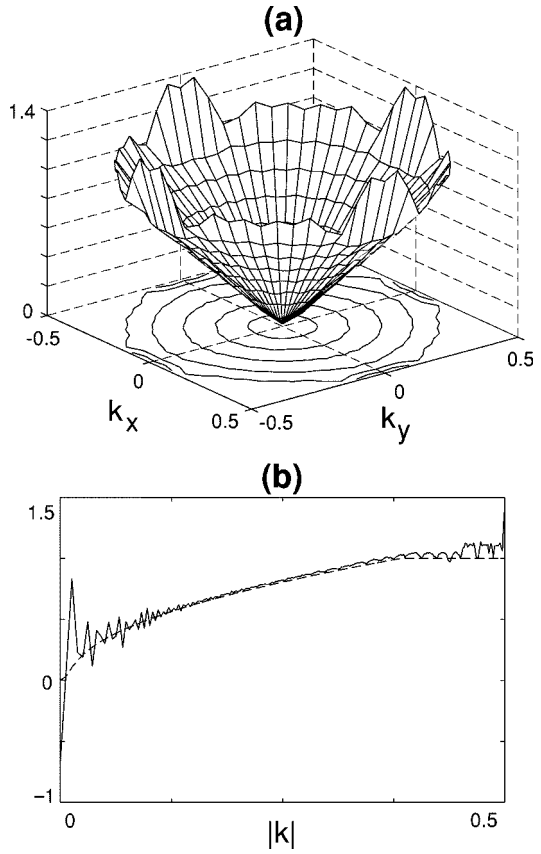


Fig. 5. The optimal compensation factors which minimize the average reconstruction error in the PSF for (a) the PR trajectory and (b) the spiral trajectory (optimal factors: solid line, conventional factors: dashed line).

Fig. 5 shows the weighted compensation factors for the PR and the spiral trajectories. Note that unlike the DCF's, which are always positive, these compensation factors can take negative values and consequently do not carry the notion of density compensation. Nevertheless, the conventional DCF's are very close to the optimal WCF's, suggesting that the reconstruction error using the conventional DCF's should be close to a minimum. To confirm this fact, we have calculated the average signal-to-error ratio (SER) in the PSF using the optimal WCF's and the conventional DCF's for both trajectories. The SER is defined as the signal amplitude, which is one for the PSF, to the root-mean-square value of the error. This definition resembles the common definition of SNR in the MR community. The SER for the PR trajectory is calculated as 30.2 dB using the conventional DCF's and 32.4 dB using the optimal WCF's. These numbers for the spiral trajectory are 29.6 and 31.8 dB, respectively.

Gridding with the weighted compensation factors is only one way of generalizing the conventional gridding algorithm. The proposed approach of the previous section allows other extensions to the gridding algorithm which otherwise may not be attainable intuitively. For instance, another way of generalization is to approximate the compensation matrix $P$ of (6) with a matrix that has a structure other than diagonal. In the simplest form, this structure should be sparse so that the resulting algorithm remains fast. An interesting structure, which is a direct extension to the conventional diagonal form, is the band

structure. With this form, the compensated measurements are obtained as the linear combination of the local samples. The approximation error with a band matrix is smaller than the error in conventional gridding with a diagonal matrix because the band structure provides extra degrees of freedom in the optimization problem. This improvement, however, comes with a price in reconstruction time which in turn restricts the maximum width of the band. The optimal band matrix can be easily obtained with the method of Appendix I. This method is studied further in Section IV in a more general form and in the context of gridding with finite kernels.

## IV. Optimal Shift-Variant Interpolating Kernel

A fast implementation of gridding reconstruction requires an interpolating kernel with a narrow width. For this reason, a $\mathrm{sinc}(\cdot)$ function, which has large side lobes, is seldom used in practice. Note that, the interpolation operation in $k$-space is equivalent to multiplication by the inverse Fourier transform of the kernel in the object domain. With a $sinc$ kernel, the image is multiplied by a rectangular function which will not affect the image inside the FOV. However, interpolating with other kernels generally results in a roll off in the intensity of the image. To compensate for this effect, the reconstructed image has to be deapodized by dividing the image by the inverse Fourier transform of the kernel. This modified form of the gridding algorithm can be expressed as

$$\hat{m}(k) = \mathcal{F}\left\{ \frac{\mathcal{F}^{-1}\left\{\sum_{i=1}^{L} h(k-\kappa_i) m(\kappa_i) d(\kappa_i)\right\}}{\mathcal{F}^{-1}\{h(k)\}} \right\} \quad (11)$$

where $\mathcal{F}\{\cdot\}$ denotes the Fourier transform operator and $h(k)$ represents the interpolating kernel.

Jackson *et al.* [13] have studied the effect of various kernels on the reconstructed image. They have shown that the Kaiser–Bessel function is a proper kernel which results in a small aliasing error. This section presents a method to find the optimal interpolating kernel and deapodization function that minimizes the gridding approximation error. We follow the same steps taken in the calculation of the optimal DCF's in the previous section. First, we rewrite (11) in a matrix form as

$$\hat{\boldsymbol{m}} = QUQ^* H^* D\mu \quad (12)$$

where $Q$ is the Fourier transform matrix and $H_{L\times N}$ is a new interpolation matrix. For a finite width kernel, the interpolation matrix has a band structure. The diagonal matrix $U_{N\times N}$ represents the deapodization matrix with the inverse Fourier transform of the kernel on its diagonal.

The optimal interpolation and deapodization matrices that minimize the gridding approximation error can be obtained by minimizing the norm of the error matrix

$$E = QUQ^* H^* D - T^* P.$$

This expression consists of the product of $H$ and $U$, thus the joint optimization of these two matrices is not a linear problem and the method of Appendix I is not directly applicable. The block diagram of Fig. 6 illustrates an iterative algorithm that

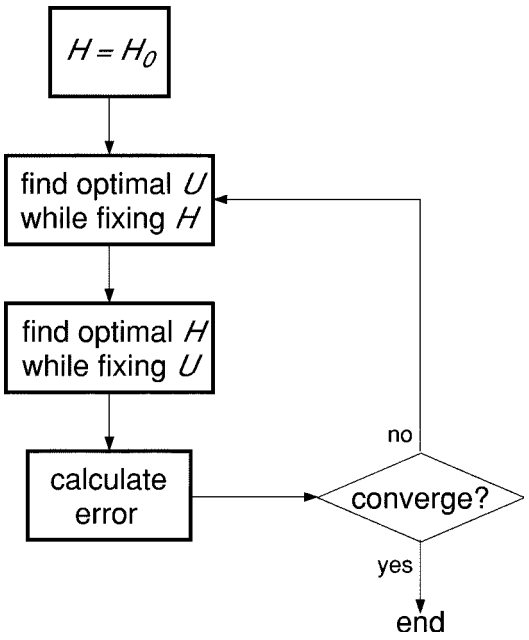


Fig. 6. An iterative algorithm for joint optimization of the interpolation and deapodization matrices.

reduces the problem to a linear problem in each iteration. The algorithm starts with an initial guess for the interpolation matrix $H$ and iterates on the following two steps until it converges to a stationary point. In the first step, the optimal deapodization matrix $U$ is calculated by fixing the interpolation matrix found in the previous step. Using this deapodization matrix, the optimal interpolation matrix is found in the second step. By assuming that one of the matrices is known in each step, the problem becomes linear in terms of the unknown matrix allowing the application of the approximation method of Appendix I. Moreover, since the norm of the error matrix reduces after each step, it is guaranteed that this method converges to a local minimum. To increase the chances of reaching the global minimum, this algorithm has to be repeated for different initial settings. The conventional values of the interpolation and deapodization matrices are found to be proper starting points in many practical cases.

It is worth noting that since the only structure imposed on the interpolation matrix is a band structure, the optimal interpolating kernel can in fact be shift variant. Therefore, although the width of the kernel is fixed, its shape can vary from one point of $k$-space to another. The intuitive approach of the conventional gridding algorithm does not support shift-variant kernels and cannot interpret the relationship between these kernels and the deapodization function.

We have applied this method to the 1-D example of Fig. 3(a). The optimal shift-variant kernels of width $3/\mathrm{FOV}$ and the corresponding deapodization function have been calculated using the proposed iterative algorithm. For the initial condition, we have used a Kaiser–Bessel function with shaping factor of $\beta = 4.2054$. This shaping factor has been found to be the best in conventional gridding for this particular kernel width [13]. The algorithm converges after only a few iterations. Fig. 7(a) shows the optimal interpolating kernel at the center of $k$-space and Fig. 7(b) shows the optimal deapodization function. The conventional interpolating kernel and the corresponding deapodiza-

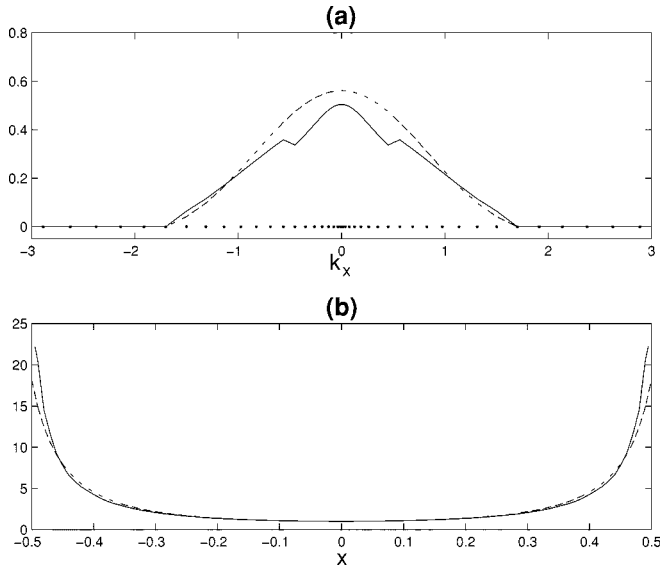


Fig. 7. (a) The optimal (solid) and conventional (dashed) interpolating kernel at the center of $k$-space. (b) The optimal (solid) and the conventional (dashed) deapodization function.

tion function are also plotted in dashed line for the sake of comparison. The average relative error of approximation using the full $sinc$ kernel is $-20.4$ dB. This error is only due to the DCF approximation and is the minimum error that one can get by tuning the interpolating kernel. With conventional gridding, this error is as high as $-14.9$ dB. Using the optimal values, the approximation error reduces to $-16.8$ dB.

## V. Fast Reconstruction from Incomplete Data

We have shown that the gridding algorithm is a fast reconstruction technique which can be derived from LSR by proper approximations. We have also proposed a method to obtain the optimal gridding parameters by minimizing the approximation error. This approach can be applied to other reconstruction techniques to reduce the complexity with a cost in the approximation error. In this section, we consider a broad range of problems in image reconstruction from incomplete data. We claim that many of these reconstruction techniques can be studied under a unified framework. We present the optimal solution for this framework and present the approximate, gridding-like solution which has reduced complexity.

Despite the recent developments in MR technology, the scan time in this imaging modality is still longer than what is desired in many applications. One approach to reduce the scan time is to acquire only an incomplete set of raw data. *A priori* knowledge about the image is to be used to fill in for the missing data. Several methods have been proposed for reconstruction from incomplete data which use various forms of prior information. For instance, in homodyne reconstruction, the prior information is the symmetry of the $k$-space data [30]. In finite support solution, the prior knowledge is on the region of the image that the object can have nonzero values [31]. The prior information on the inherent local spatial and temporal resolution have also been used in locally focused MRI [32], [33] and partial FOV reconstruction [34]–[36]. Some other methods of reconstruction from incomplete data are: keyhole imaging and UNFOLD that use prior knowledge about the temporal behavior in the spatial frequency domain [37]–[39] and feature-recognizing MRI that uses the salient features derived from similar images [40]. The practical use of some of these reconstruction techniques is hindered partially because of their often complex reconstruction algorithm.

In many of these techniques, the prior information can be represented as an affine constraint on the image in the following form

$$\boldsymbol{m} = C\boldsymbol{x} + \boldsymbol{m}_0, \tag{13}$$

where $\boldsymbol{m}$ is the vector of grid points to be reconstructed, $\boldsymbol{m}_0$ is a vector that is known in advance, and $\boldsymbol{x}$ is an unknown vector. $C_{N\times P}$ is the constraint matrix with a rank of $P < N$. This matrix reduces the degrees of freedom of the problem from $N$ to $P$, thus allowing an incomplete set of acquired data to represent the image uniquely. For instance, in partial FOV reconstruction where a portion of the image is assumed to be known in advance, $\boldsymbol{m}_0$ represents the Fourier transform of the known part and $\boldsymbol{x}$ is the pixel values of the unknown section and $C$ is the matrix that maps $\boldsymbol{x}$ to a full FOV image and calculates its Fourier transform.

The combination of (3) and (13) forms a framework that can express many reconstruction techniques. The minimum-squared-error solution under this framework is expressed as

$$\hat{\boldsymbol{m}} = (C^*T^*TC)^{-1}C^*T^*(\boldsymbol{\mu} - T\boldsymbol{m}_0). \tag{14}$$

This solution, like LSR, suffers from high complexity. A fast implementation of this algorithm may be obtained by approximating some of the involved matrices with sparse (or other properly structured) matrices. This approximation can be done in many different ways with respect to the matrices being approximated or the structures being used.

To illustrate this approach, an example for a case of locally focused MRI using the 1-D variable-density trajectory of Fig. 3(a) is studied. The image is assumed to have twice the original FOV. Therefore, the sampling density at the periphery of $k$-space is half of the Nyquist requirement, resulting an incomplete set of raw data. It is, however, assumed that the inherent resolution of the object outside the original FOV is only $1/5$ of the resolution inside (Fig. 8). The $k$-space trajectory of Fig. 3(a) provides a proper sampling pattern for this particular object model. Note that Fig. 3(b), which is a rough plot of the local sampling period, indicates that the center of $k$-space is sampled dense enough for an aliasing-free reconstruction of the whole image. This central area covers $1/5$ of $k$-space, providing low–resolution information for the large FOV. The rest of raw data, which consists of the low-density samples of the high spatial frequency content of the image, provides high resolution information for only half of the whole FOV.

This image model can be expressed in the affine form of (13) with $\boldsymbol{m}_0 = 0$. The constraint matrix $C$ can be obtained using the method described in [32] or simply by collecting the desired independent PSF's as the columns of this matrix. In its simplest form, the PSF at the low resolution section of the image can be represented by a rectangular function with a width of 5 pixels. A gridding-like reconstruction for this example is carried out in

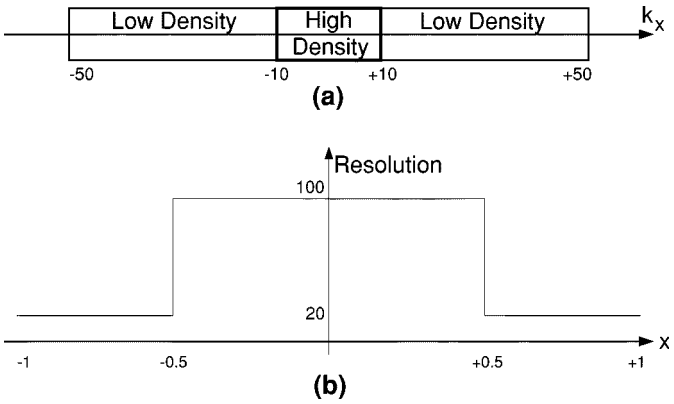


Fig. 8. (a) The 1-D variable-density trajectory provides high sampling density ($>2 \times \mathrm{FOV}$) at the center of $k$-space and low density ($>1 \times \mathrm{FOV}$) at the outer parts. (b) The extent of the object is twice the designed FOV. The inherent resolution of the object in the outer section is only $1/5$ of the resolution in the central part. (All numbers are normalized to the designed FOV.)

two steps. First, the central section of $k$-space is gridded over a dense grid to form the low-resolution information for the whole FOV. Then the outer part of the $k$-space is gridded over a sparse grid to provide the high-resolution data for the central part of the image. These two sections are added together to form the final reconstructed image. These operations can be expressed in the following form:

$$\hat{\boldsymbol{m}} = Q_h U_h Q_h^* H_h D_h \boldsymbol{\mu}_h + Q_l U_l Q_l^* H_l D_l \boldsymbol{\mu}_l. \tag{15}$$

$\boldsymbol{\mu}_h$ ($\boldsymbol{\mu}_l$) denotes the raw data corresponding to the high (low) density central (outer) part of the $k$-space. $D_h$ ($D_l$) is the diagonal density compensation matrix for the central (outer) $k$-space points. $H_h$ ($H_l$) is the interpolation matrix for the dense (sparse) grid. $U_h$ ($U_l$) represents the deapodization matrix for the low- (high-) resolution part of the image. $Q_h$ ($Q_l$) is the Fourier transform matrix that maps the low- (high-) resolution image to the dense (sparse) grid. The interpolation matrices have band structures. A larger kernel width is used for $H_l$ which interpolates the sparse grid. The deapodization matrices have a diagonal structure.

To obtain the optimal interpolation and deapodization matrices, we rewrite (15) in a more familiar form of (12) as

$$\hat{\boldsymbol{m}} = [Q_h \quad Q_l] \begin{bmatrix} U_h & 0 \\ 0 & U_l \end{bmatrix} \begin{bmatrix} Q_h^* & 0 \\ 0 & Q_l^* \end{bmatrix} \times \begin{bmatrix} H_h & 0 \\ 0 & H_l \end{bmatrix} \begin{bmatrix} D_h & 0 \\ 0 & D_l \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_h \\ \boldsymbol{\mu}_l \end{bmatrix}.$$

Note that the second and forth augmented matrices are unknowns to be found using the iterative algorithm of the previous section.

We have calculated the error using interpolating kernels with widths of three times their corresponding grid units. The average relative approximation error using the Kaiser–Bessel function with shaping factor of $\beta = 4.2054$ is $-12.3$ dB. This error is $-15.8$ dB when the optimal gridding parameters are used. Fig. 9 shows the optimal and the conventional interpolating kernels at the center of $k$-space.

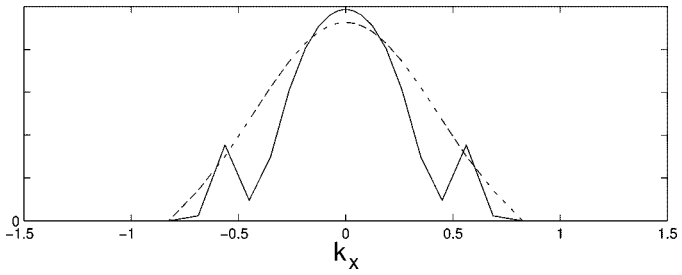


Fig. 9. The optimal (solid) and the conventional (dashed) interpolating kernels for the simple case of locally-focused reconstruction using a variable density trajectory.

It is worth noting that, the frame-work introduced in this section can be extended to incorporate some prior statistical information on the image, similar to the method described in [41] and [42]. A particular case of interest is when the second-order statistics of the image is available in advance. In this case, one can approximate each pixel as a linear combination of the other pixels such that the average error of approximation is minimal. In other words, knowing the second-order statistics, one can derive a matrix $R$ from the autocorrelation function of the image such that $R\boldsymbol{m}$ represents the approximation error vector with minimum norm. This requirement can be forced to the least squares solution by using the following augmented form of (3):

$$\begin{bmatrix} \boldsymbol{\mu} \\ 0 \end{bmatrix} = \begin{bmatrix} T \\ \sigma R \end{bmatrix} \boldsymbol{m}. \tag{16}$$

The least squares solution to this equation minimizes the norm of

$$\|\boldsymbol{\mu} - T\boldsymbol{m}\|^2 + \sigma^2 \|R\boldsymbol{m}\|^2.$$

The first term, by itself, yields the solution of (4) and the second term imposes the desired second-order statistics on the reconstructed image. $\sigma$ is a weighting factor that adjusts the proper emphasis on these two terms. It can be shown that this solution generates an image with maximum *a posteriori* probability when the image and noise have Gaussian distributions. Note that , the combination of equations (16) and (13) can be used to impose both the statistical and affine constraints on the reconstructed image. The resulting least squares solution may be properly approximated for fast implementation.

## VI. Applications to 2-D Spiral MR Imaging

Data acquisition along a spiral trajectory [16], [43] offers many benefits in magnetic resonance imaging. Spiral trajectory provides a more efficient coverage of $k$-space, resulting in a shorter scan time. This type of acquisition also exhibits more immunity to flow and motion artifacts. Gridding is the most prevailing reconstruction technique for spiral acquisition. In this section, we calculate the optimal gridding parameters for the spiral trajectory of Fig. 1(c) and compare the average reconstruction error in an exemplary 2-D image using the optimal and the conventional parameters.

The phantom image of Fig. 10(a) is used as a reference image. The spiral raw data for this image is synthetically generated by taking the Fourier transform of the image to obtain the grid

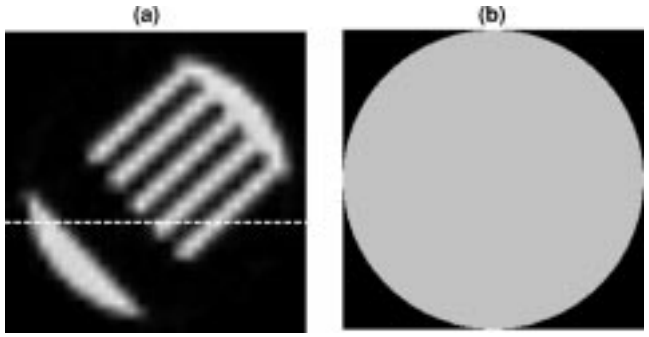


Fig. 10. (a) The reference image to be used in calculation of the reconstruction error. (b) The support region of the reference image to be used as the prior information for reconstruction from incomplete data.

points and then interpolating the spiral samples from the grid data using a $\mathrm{sinc}(\cdot)$ kernel. The computational complexity and the memory requirement of our algorithm are on the order of $N^2$ where $N$ is the grid size. We limit the grid size to $64 \times 64$ to avoid excessive complexity. Note that, this complexity is only related to the calculation of the gridding parameters. This calculation is done only once when a new trajectory is designed. The image reconstruction algorithm, however, has the same order of complexity as the conventional gridding reconstruction.

We calculate the optimal shift-variant interpolating kernel of width $2/\mathrm{FOV}$ such that a weighted average of the approximation error is minimum. The weighting matrix is chosen to minimize the reconstruction error in the PSF. We solve the problem as a case of reconstruction from incomplete data using information on the support region of the image as the prior knowledge [31]. We assume that object is confined within the circle inscribed in the rectangular FOV and everything outside this circular region have zero values. This support region is shown in Fig. 10(b). This prior knowledge can be modeled as

$$\boldsymbol{m} = QR\boldsymbol{x}$$

where $\boldsymbol{x}_P$ is a vector consisting of all pixel values within the circular FOV, $Q_{N\times N}$ is the Fourier transform matrix that maps the pixel values of an image with rectangular FOV to $k$-space grid point, and $R_{N\times P}$ $(N > P)$ is a matrix that adds zero-value pixels to $\boldsymbol{x}$ such that the resulting vector represents an image with rectangular FOV. $R$ is obtained by extracting the columns of an $N \times N$ identity matrix that correspond to the pixels outside the circular support region.

The norm of the following approximation error matrix is minimized:

$$E = W_l Q^* U Q H W_r - I_{P\times P}$$

where $W_l = R^*$ and $W_r = TQR$ are the left and right weighting matrices that effectively yield a minimum reconstruction error in the point-spread function inside the circular support region. $H$ and $U$ are the unknown interpolating and deapodization matrices to be calculated using the iterative algorithm of Section IV. Note that, since we are considering a shift-variant kernel we are able to combine the interpolating and compensation matrices into a single matrix denoted by $H$.

The image is reconstructed from the synthetic raw data using both the optimal and the conventional gridding parameters. For the conventional gridding parameters, we use the DCF's suggested in [16] and a Kaiser–Bessel function with the shaping factor of $\beta = 2.3934$ as the interpolating kernel. This shaping factor has been found to be the best for the corresponding kernel width [13]. The reconstruction error, calculated by subtracting the reconstructed images from the reference image, is shown in Fig. 11. Note that these images are scaled by a factor of four for a better depiction of the error. Fig. 11(c) shows the error for the high lighted line of Fig. 10(a). The results show that for almost all pixels the reconstruction error is smaller using the optimal gridding parameters. The average signal to reconstruction error ratio over the entire image is 21.14 dB using the conventional gridding parameters and 25.93 dB using the optimal parameters, an enhancement of more than 4 dB.

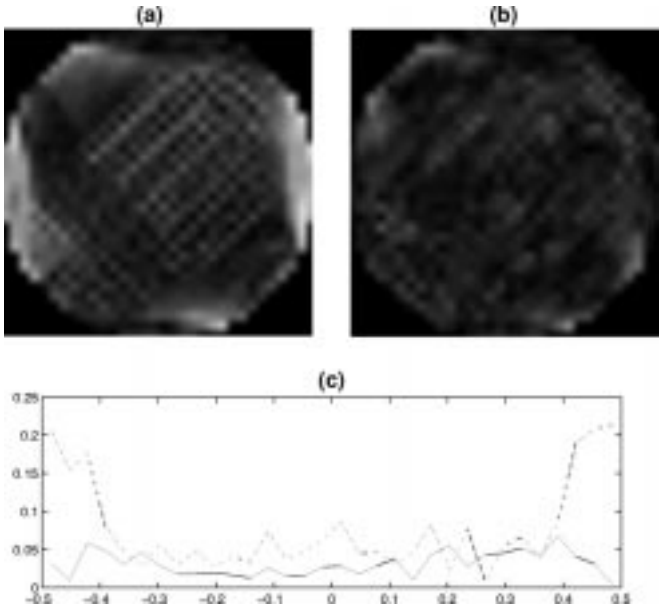


Fig. 11. The magnitude of the reconstruction error using (a) the conventional gridding parameters and (b) the optimal gridding parameters (the images are scaled by a factor of four for better visualization of the error). (c) The reconstruction error over the high lighted line of Fig. 10 using the optimal (solid) and the conventional (dashed) gridding parameters.

Note that, in the analysis presented in this paper, we have ignored the effect of measurement noise for simplicity. However, it is not difficult to incorporate the effect of noise in the least squares solution. The optimal gridding parameters that take into account the measurement noise are obtained with a proper approximation of this least squares solution.

## VII. Summary

We showed that the gridding reconstruction algorithm is an approximation to the least-squares solution. The optimal gridding parameters minimize the approximation error. These parameters are calculated by minimizing the Frobenious norm of an approximation error matrix. We presented a method to solve this problem in a more general form of approximation using linearly structured matrices. Based on this approach, a closed form solution for the optimal DCF's is obtained. We showed that these factors are in fact inversely related to the local sampling density, thus carrying the notion of density compensation.

The optimal interpolating kernel and the optimal deapodization function are more difficult to calculate because the joint optimization of these two factors is a nonlinear problem. We proposed an iterative algorithm that reduces this nonlinear problem to a linear one in each step, allowing the application of our method of matrix approximation. The algorithm is guaranteed to converge to a local minimum and our simulation shows that the rate of convergence is rather fast.

Our method of calculating the gridding parameters provides solutions for situations that intuition may not be helpful and the notion of density compensation or the Fourier relationship between the interpolating kernel and the deapodization function are not valid anymore. We studied examples of such cases in gridding with shift-variant kernels and reconstruction targeted for minimum error in the PSF. We also illustrated the utility of our method in developing gridding-like algorithms for fast implementation of methods of reconstruction from incomplete data.

We applied this technique to find the optimal gridding parameters for a case of 2D spiral MR imaging. We calculated the shift-variant kernel that yields the minimum error in the PSF. The problem was solved as a case of reconstruction from incomplete data where the prior information was the support region of the image. We obtained more than 4-dB reduction in the reconstruction error using the optimal gridding parameters.

Our method of calculating the optimal gridding parameters is computationally demanding. This problem can be alleviated by taking advantage of the sparse structure of many matrices involved. However, note that, this complexity needs to be tolerated only when a new sampling trajectory is designed. Once the optimal parameters are calculated, the complexity of the reconstruction algorithm is still as low as the complexity of the conventional gridding reconstruction algorithm.

## Appendix I
## Optimal Structured Matrix Approximation

In this section, we study the general problem of matrix approximation with structured matrices. We specifically consider the so-called linearly structured matrices. A matrix has a linear structure if all of its elements can be written as a linear combination of a few free variables. Many well-known structures, such as diagonal, triangular, band, sparse, Toeplitz, Hankel, etc., can be categorized as special cases of the linear structure. A matrix $X$ has a linear structure if and only if

$$\mathrm{vec}\{X\} = S\boldsymbol{v} \tag{17}$$

where the $\mathrm{vec}\{\cdot\}$ operator maps $X$ to a tall vector obtained by stacking all the columns of $X$ on top of each other. $\boldsymbol{v}$ represents the vector of free variables and $S$ is a matrix that enforces a specific linear structure (such as diagonal or Toeplitz) on $X$. For instance, if $X$ is a $3 \times 3$ symmetric Toeplitz matrix, then $\boldsymbol{v}$ is the first column of $X$ and

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}'.$$

Similarly, if $X$ is a $3 \times 3$ diagonal matrix, then $\mathbf{v}$ consists of the diagonal elements of $X$ and $S$ is obtained by inserting three rows of zeros under the first two rows of the $3\times3$ identity matrix.

With this introduction, we continue with the matrix approximation problem:

*Problem Statement:* Given matrices $A_{p\times m}, B_{n\times q}$ and $C_{p\times q}$, find a matrix $X_{m\times n}$ such that:

1) $X$ has a predefined linear structure;
2) the product $AXB$ is a good approximation to matrix $C$.

The approximation error matrix is defined as $E = AXB - C$. The Frobenious norm of this matrix, $\|E\|_F$, is chosen as a measure of appropriateness of the approximation. The unknown matrix $X$ is calculated such that this norm is minimum. The choice of Frobenious norm from the variety of other matrix norms, such as 2-norm, is due to two reasons. First, the Frobenious norm provides a measure of average induced error which may be more appropriate in many applications because it is a less conservative indicator than the 2-norm which is a measure of the maximum induced error. Second, with the Frobenious norm, the solution to this problem is mathematically tractable which may not be the case with other norms.

Note that if no particular structure were imposed on $X$, the solution would be easily obtained by multiplying $C$ from left and right by the pseudo-inverses of $A$ and $B$, respectively. The complexity of the problem is mainly due to the fact that the solution must have the predefined linear structure. We start the solution with the following two lemmas from basic matrix theory.

*Lemma 1:* The Frobenious norm of a matrix is equal to the squared length of the $\mathrm{vec}\{\cdot\}$ of that matrix

$$\|A\|_F = |\mathrm{vec}\{A\}|^2.$$

*Lemma 2:* For matrices $A_{p\times m}, X_{m\times n}$ and $B_{n\times q}$

$$\mathrm{vec}\{AXB\} = (B' \otimes A)\,\mathrm{vec}\{X\}$$

where $B'$ denotes the transpose (and not the conjugate transpose) of $B$ and $(B' \otimes A)$ represents the Kronecker (or direct) product of $B'$ and $A$ [44].

Using these two lemmas, we can express the norm of error matrix as

$$\begin{aligned} \|AXB - C\|_F &= |\mathrm{vec}\{AXB - C\}|^2 \\ &= |\mathrm{vec}\{AXB\} - \mathrm{vec}\{C\}|^2 \\ &= |(B' \otimes A)\,\mathrm{vec}\{X\} - \mathrm{vec}\{C\}|^2. \end{aligned}$$

Now the problem is to find a matrix $X$ such that the length of vector $|(B' \otimes A)\,\mathrm{vec}\{X\} - \mathrm{vec}\{C\}|$ is minimum. This problem is equivalent to finding the least squares solution to the following linear equation:

$$(B' \otimes A)\,\mathrm{vec}\{X\} = \mathrm{vec}\{C\}$$

where $\mathrm{vec}\{X\}$ is the unknown vector. The next step is to solve this least squares problem under the constraint of a linear structure. To enforce the desire structure on $X$, we use (17) to substitute for $X$ in the previous equation and to obtain

$$(B' \otimes A)S\mathbf{v} = \mathrm{vec}\{C\}. \tag{18}$$

The final solution to the matrix approximation problem is obtained from (17) where $\boldsymbol{v}$ is the least squares solution to (18).

## REFERENCES


[1] W. N. Brouw, "Apperture synthesis," in *Methods in Computational Physics*, B. Alder, S. Fernbach, and M. Rotenberg, Eds. New York: Academic, 1975, pp. 131–175.

[2] G. T. Herman, *Image Reconstruction from Projections: Implementation and Applicaitons*. Berlin, Germany: Springer-Verlag, 1979.

[3] R. K. Mueller, M. Kaveh, and G. Wade, "Reconstructive tomography and applications to ultrasonics," *Proc. IEEE*, vol. 67, pp. 567–587, 1979.

[4] K. F. King and P. R. Moran, "A unified description of NMR imaging, data collection strategies and reconstruction," *Magn. Reson. Med.*, vol. 11, pp. 1–14, 1984.

[5] R. N. Bracewell and A. C. Riddle, "Inversion of fan-beam scans in radio astronomy," *Astrophys. J.*, vol. 150, no. 2, pp. 424–437, 1967.

[6] A. J. Devaney, "A filtered backpropagation algorithm for diffraction tomorgaphy," *Ultrason. Imag.*, no. 4, pp. 336–350, 1982.

[7] S. X. Pan and A. C. Kak, "A computational study of reconstruction algorithms for diffraction tomography: Interpolation versus filtered backpropagation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 1262–1275, 1983.

[8] H. Stark, J. W. Woods, I. Paul, and R. Hingorani, "Direct Fourier reconstruction in computer tomography," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, pp. 237–245, 1981.

[9] ——, "An investigation of computerized tomography by direct Fourier inversion and optimum interpolation," *IEEE Trans. Biomed. Eng.*, vol. 28, pp. 496–505, 1981.

[10] D. E. Hogg, G. H. MacDonald, R. G. Conway, and C. M. Wade, "Synthesis of brightness distribution in radio sources," *Astronom. J.*, vol. 74, no. 10, pp. 1206–1213, 1969.

[11] A. R. Thompson and R. N. Bracewell, "Interpolation and Fourier transformation of fringe visibilities," *Astronom. J.*, vol. 79, no. 1, pp. 11–24, 1974.

[12] J. D. O'Sullivan, "A fast sinc function gridding algorithm for Fourier inversion in computer tomorgaphy," *IEEE Trans. Med. Imag.*, vol. MI-4, pp. 200–207, 1985.

[13] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski, "Selection of a convolution function for Fourier inversion using gridding," *IEEE Trans. Med. Imag.*, vol. 10, pp. 473–478, 1991.

[14] M. Soumekh, "Reconstruction and sampling constraints for spiral data," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 37, pp. 882–891, 1989.

[15] C. J. Hardy, H. E. Cline, and P. A. Bottomly, "Correcting for nonuniform $k$-space sampling in two-dimensional NMR selective excitation," *J. Magn. Reson.*, vol. 87, pp. 639–645, 1990.

[16] C. H. Meyer, B. S. Hu, D. G. Nishimura, and A. Macovski, "Fast spiral coronary artery imaging," *Magn. Reson. Med.*, vol. 28, pp. 202–213, 1992.

[17] O. Heid, "Archimedian spirals with Euclidean gradient limits," in *Proc. ISMRM, 4th Annu. Meeting*, 1996, p. 114.

[18] R. D. Hoge, R. K. S. Kwan, and G. B. Pike, "Density compensation functions for spiral MRI," *Magn. Reson. Med.*, vol. 38, pp. 117–128, 1997.

[19] J. Liao and N. Pelc, "Image reconstruction of generalized spiral trajectory," in *Proc. ISMRM, 4th Annu. Meeting*, 1996, p. 359.

[20] V. Rasche and R. Proksa, "Reconstruction of MR images from data sampled along arbitrary $k$-space trajectory," in *Proc. ISMRM, 6th Annu. Meeting*, 1998, p. 668.

[21] A. M. Takahashi, "Consideration for using the Voronoi areas as a $k$-space weighting function," in *Proc. ISMRM, 7th Annu. Meeting*, 1999, p. 92.

[22] J. G. Pipe and P. Menon, "Sampling density compensation in MRI: Rationale and an iterative numerical solution," *Magn. Reson. Med.*, vol. 41, pp. 179–186, 1999.

[23] J. H. Zwaga, F. T. A. W. Wajer, R. de Beer, M. Fuderer, A. F. Mehlkopf, and D. van Ormondt, "Improved Kaiser-Besel window parameter selection for gridding," in *Proc. ISMRM, 6th Annu. Meeting*, 1998, p. 669.

[24] D. Chen, R. B. Marr, and P. C. Lauterbur, "Reconstruction from NMR data acquired with imaging gradinets having arbitrary time dependence," *IEEE Trans. Med. Imag.*, vol. MI-5, pp. 162–164, 1986.

[25] D. Rosenfeld, "An optimal and efficient new gridding algorithm using singular value decomposition," *Magn. Reson. Med.*, vol. 40, pp. 14–23, 1998.

[26] D. Twieg, "The $k$-trajectory formulation of the NMR imaging process with application in analysis and synthesis of imaging methods," *Med. Phys.*, vol. 10, pp. 610–621, 1983.

[27] H. Schomberg and J. Timmer, "The gridding method for image reconstruction by Fourier transformation," *IEEE Trans. Med. Imag.*, vol. MI-14, pp. 596–607, 1995.

[28] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Press, 1984.

[29] G. Herman, *Image Reconstruction from Projections*. New York: Academic, 1980.

[30] D. C. Noll, D. G. Nishimura, and A. Macovski, "Homodyne detection in magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 10, no. 2, pp. 154–163, 1991.

[31] S. K. Plevritis and A. Macovski, "Spectral extrapolation of spatially bounded images," *IEEE Trans. Med. Imag.*, vol. 14, no. 3, pp. 487–497, 1995.

[32] Y. Cao, D. N. Levin, and L. Yao, "Locally-focued MRI," *Magn. Reson. Med.*, vol. 34, pp. 858–867, 1995.

[33] L. Yao, Y. Cao, and D. N. Levin, "2d locally focued MRI: Applications to dynamic and spectroscopic imaging," *Magn. Reson. Med.*, vol. 36, pp. 834–846, 1996.

[34] X. Hu and T. Parrish, "Reduction of field of view for dynamic imaging," *Magn. Reson. Med.*, vol. 31, pp. 691–694, 1994.

[35] K. Scheffler and J. Hennig, "Reduced circular field-of-view imaging," *Magn. Reson. Med.*, vol. 40, pp. 474–480, 1998.

[36] H. Sedarat, A. B. Kerr, J. M. Pauly, and D. G. Nishimura, "Partial-FOV reconstruction in dynamic spiral imaging," *Magn. Reson. Med.*, vol. 43, pp. 429–439, 2000.

[37] J. J. van Valls, M. E. Brummer, W. T. Dixon, H. H. Tuithof, H. Engels, R. C. Nelson, B. M. Gerety, J. L. Chezmar, and J. A. den Boer, "'Keyhole' method for accelerating imaging of contrast agent uptake," *J. Magn. Reson. Imag.*, vol. 3, pp. 671–675, 1993.

[38] R. A. Jones, O. Haraldseth, T. B. Muller, P. A. Rinck, and A. N. Oksendal, "$K$-space substitution: A novel dynamic imaging technique," *Magn. Reson. Med.*, vol. 29, pp. 830–834, 1993.

[39] B. Madore, G. H. Glover, and N. J. Pelc, "UNaliasing by Fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fmri," in *Proc. ISMRM, 6th Annu. Meeting*, 1998, p. 575.

[40] Y. Cao and D. N. Levin, "Feature-recognizing MRI," *Magn. Reson. Med.*, vol. 30, pp. 305–317, 1993.

[41] F. T. A. W. Wajer, R. de Beer, M. Fuderer, A. F. Mehlkopf, and D. van Ormondt, "Spiral MRI scan-time reduction through omission of interleaves and Bayesian image reconstruction," in *Proc. ISMRM, 6th Annu. Meeting*, 1998, p. 1960.

[42] ——, "Bayesian image reconstruction from an arbitrary and sparsely sampled $k$-space without gridding and density correction," in *Proc. ISMRM, 6th Annu. Meeting*, 1998, p. 667.

[43] C. B. Ahn, H. H. Kim, and Z. H. Cho, "High-speed spiral-scan echo planar NMR imaging," *IEEE Trans. Med. Imag.*, vol. MI-5, pp. 2–7, 1986.

[44] P. Lancaster and M. Tismenetsky, *Theory of Matrices*. Orlando, FL: Academic, 1985.