This chapter provides a tutorial overview of advanced image reconstruction methods used in MRI. The term “advanced” is used loosely to refer to the class of non-Fourier reconstruction methods developed for handling the inverse problem with limited Fourier samples. We will consider two specific cases: (a) the superresolution reconstruction problem (associated with limited Fourier samples collected at the Nyquist rate) and (b) the parallel imaging problem (arising when Fourier samples are collected at sub-Nyquist rates, using multiple nonuniform receiver channels).

For notational convenience, we will consider only the one-dimensional case. The following is a summary of notations used in this chapter.

\[ \rho(x) \quad \text{Desired image} \]
\[ \hat{\rho}(x) \quad \text{Reconstructed image} \]
Advanced Image Processing in Magnetic Resonance Imaging

\[ D(k_n) = \int_{-W/2}^{W/2} \rho(x)s(x)e^{-2\pi ik_n x} dx, \]

where we explicitly include the sensitivity weighting function \( s(x) \) of the receiver coil. In conventional Fourier imaging, \( s(x) \) is often ignored because it can be assumed to be a constant over the field of view (FOV), and \( D(k_n) \) is usually measured at \( k_n = n\Delta k \) for \( n = -N/2, -N/2 + 1, \ldots, N/2 - 1 \), with \( N \) being the total number of encodings acquired. In multichannel Fourier imaging (often known as parallel imaging), an array of receiver channels (or coils) with sensitivity functions \( s_\ell(x) \) is used to acquire \( D_\ell(k_n) \) simultaneously for \( \ell = 1, 2, \ldots, L \). To increase imaging speed, \( D_\ell(k_n) \) is measured at \( k_n = n\Delta k \) for \( n = -M/2, -N/2 + 1, \ldots, M/2 - 1 \), with \( M = N/R \) and \( \Delta k = R\Delta k \). In other words, the \( k \)-space signal is measured at a sub-Nyquist rate in each receiver channel. Before we discuss advanced techniques to handle the image reconstruction problem associated with these two data acquisition schemes, a brief review of the popular Fourier reconstruction method is in order.

## 2.2 FOURIER RECONSTRUCTION

Given \( D(n\Delta k) \) and \( s(x) = 1 \), it is well known that \( \rho(x) \) can be reconstructed using the Fourier series, that is,

\[ \rho(x) = \Delta k \sum_{n=-\infty}^{\infty} D(n\Delta k)e^{2\pi ik_n x}, \quad |x| < \frac{1}{2\Delta k}. \]
For finite sampling, there are not sufficient data to define this series. The conventional Fourier reconstruction method treats the unknown coefficients as zero and, as a result, we have

\[ \hat{\rho}(x) = \Delta k \sum_{n=-N/2}^{N/2-1} D(n\Delta k) e^{i2\pi n\Delta x}, \quad |x| < \frac{1}{2\Delta k}, \]  

(2.3)

which can be evaluated efficiently using the fast Fourier transform (FFT) algorithm. Some basic properties of the Fourier reconstruction method is summarized in the following remarks:

**Remark 1:** Given \( D(n\Delta k) \) for \( n = -N/2, -N/2 + 1, \ldots, N/2 - 1 \), any \( \hat{\rho}(x) \) given below satisfies Equation 2.1,

\[ \hat{\rho}(x) = \Delta k \sum_{n=-N/2}^{N/2-1} D(n\Delta k) e^{i2\pi n\Delta x} + \sum_{n=-N/2, n \neq N/2} c_n e^{i2\pi n\Delta x}, \]  

(2.4)

which is often called a feasible reconstruction of \( \rho(x) \).

**Remark 2:** The Fourier reconstruction, \( \hat{\rho}(x) \), given in Equation 2.3 is the minimum-norm feasible solution because

\[ \hat{\rho}(x) = \frac{1}{2\Delta x} \arg \min_{\rho} \int_{\Delta x}^{\Delta x} |\hat{\rho}(x)|^2 dx = 0. \]  

(2.5)

**Remark 3:** The Fourier reconstruction, \( \hat{\rho}(x) \), is related to the true image \( \rho(x) \) by

\[ \hat{\rho}(x) = \int_{\Delta x}^{\Delta x} \rho(\hat{x}) h(x - \hat{x}) d\hat{x}, \]  

(2.6)

where \( h(x) \), known as the point spread function (PSF), is given by

\[ h(x) = \Delta k \frac{\sin(\pi N\Delta k x)}{\sin(\pi \Delta k x)} e^{-i\pi \Delta k x}. \]  

(2.7)

Note that \( h(x) \) is a periodic function, and within each period it is similar to a sinc function. The width of its main lobe, as measured by the interval between the first two zero crossings, is \( 2/(N\Delta k) \). The effective width \( W_h \) of \( h(x) \) is often
taken to be the width of an approximating rectangular pulse with height \(h(0)\) and the same area. It is easy to show that

\[
W_x = \frac{1}{h(0)} \int_{\frac{1}{N\Delta k}}^{1} h(x)dx = \frac{1}{N\Delta k},
\] (2.8)

which is exactly half the width of the main lobe of \(h(x)\).

The right-hand side of Equation 2.8 is known as the Fourier pixel size, in contrast to the usual image pixel size \(\Delta x\). Note that \(\Delta x\) can be made arbitrarily small using any signal interpolation schemes, but image resolution is fundamentally limited to \(1/(N\Delta k)\). Another implication of Equation 2.8 is that \(W_x\) and \(N\) cannot be reduced simultaneously; in other words, improving image resolution and reducing the number of measured data points cannot be achieved simultaneously.

In addition to a loss of resolution in \(\hat{\rho}(x)\), the convolution operation in Equation 2.6 also results in the well-known Gibbs ringing artifact in \(\hat{\rho}(x)\). This artifact manifests itself as spurious ringing around sharp edges, as illustrated in Figure 2.1. The maximum undershoot or overshoot of the spurious ringing is about 9% of the intensity discontinuity and is independent of the number of data points used in the reconstruction. The frequency of oscillation, however, increases as more data points are used. For this reason, when a large number of data points is used in practice, the spurious ringing does not cover an appreciable distance in the reconstructed image and thus becomes invisible.

**FIGURE 2.1** Gibbs ringing artifacts.
2.3 CONSTRAINED IMAGE RECONSTRUCTION

For years, the belief existed that information beyond the measurement cutoff frequency was not recoverable, thus the Rayleigh resolution limit [1]. Although the information is not apparent in the measured data, we now have learned how to look elsewhere for the additional information required to restore those frequency contents [2,3]. Constrained methods are the mathematical tools developed to accomplish this objective by using a priori information to compensate for the lack of high-frequency experimental data in the reconstruction process. Although constrained data processing methods have been used extensively for decades in other fields, application of the constrained reconstruction concept to MRI is very recent. The first successful effort was perhaps due to Smith [4] and, since then, research interest in this area has continued to grow for at least two reasons: first, the rapid development of computing technology has made it possible to use computation-intensive algorithms for practical applications and, second, the advantages of modern constrained reconstruction methods have made them worthwhile. In particular, the ability to reduce data truncation artifacts and improve image resolution is very desirable and can produce effects unmatched by the traditional unconstrained Fourier methods. Nonparametric constraints permit the use of the conventional Fourier series model for image function, and reconstruction methods of this type usually involve explicit data extrapolation to recover some of the unmeasured (presumably lost) high-spatial-frequency data so as to reduce truncation artifact. Parametric modeling methods, on the other hand, represent the image function in terms of a set of parameterized basis functions, rather than the nonparameterized harmonic sinusoidal functions used in the Fourier series. These methods can, in principle, generate images of infinite resolution from the model without explicitly extrapolating the data to the infinite frequency range. In this sense, parametric model constraints are often more powerful than nonparametric constraints, although sometimes they may not be as robust. Explicit data extrapolation is also possible and often used with parametric methods by using the model to generate the unmeasured data.

2.3.1 NONPARAMETRIC METHODS

A popular mathematical algorithm used in many nonparametric reconstruction methods is alternate projection, or projection onto convex sets (POCS). The principle of POCS has been discussed in great detail in the signal processing literature. We review here only the central ideas and give a couple of examples of its use in MRI.

**Definition:** A subset $\Omega$ in the Hilbert space $H$ is said to be convex if together with any $x_1$ and $x_2$, it also contains $\mu x_1 + (1 - \mu)x_2$ for all $\mu, 0 \leq \mu \leq 1$.

**Definition:** For any $x \in H$, the projection $P_\Omega x$ of $x$ onto $\Omega$ is the element in $\Omega$ closest to $x$. If $\Omega$ is closed and convex, $P_\Omega$ exists and is uniquely determined by $x$ and $\Omega$ from the following minimality criterion

$$\| x - P_\Omega x \| = \min_{y \in \Omega} \| x - y \|.$$
Briefly, the method of POCS is simply an iterative algorithm that finds a solution subject to a number of convex-type constraints by alternate projection. The following theorem is central to this technique which assures the convergence of such an iterative process.

**Theorem:** Given $m$ closed convex sets $\Omega_i$, $i = 1, 2, \ldots, m$, in $H$ and their corresponding projection operators $P_i$, if $\Omega_0 = \bigcap_{i=1}^m \Omega_i$ is nonempty, the sequence generated by

$$f_{l+1} = P_m P_{m-1} \ldots P_1 f_l, \quad l = 0, 1, \ldots$$

converges (weakly) to an element $f \in \Omega_0$ for any initial value $f_0 \in H$.

For problems with $m$ pieces of *a priori* constraints of which each restricts the solution to a convex set, POCS is an ideal method for finding a solution. Various type of convex-type constraints exist for the reconstruction problem addressed. For example, the following constraints are of convex type:

1. Data-consistency constraint:

   $$\Omega_1 = \{ \rho(x) : F(\rho(x)) = D_k, \text{ for all the measured } k \text{ values} \}.$$

2. Limited-support constraint:

   $$\Omega_2 = \{ \rho(x) : \rho(x) = 0, \quad \text{for } |x| > W/2 \}.$$

3. Bounded-magnitude constraint:

   $$\Omega_3 = \{ \rho(x) : |\rho(x)| \leq B \}.$$

4. Phase constraint:

   $$\Omega_4 = \{ \rho(x) : \text{arg}(\rho(x)) = \theta(x) \}.$$

5. Bounded-noise variance constraint:

   $$\Omega_5 = \{ \rho(x) : \|s - F(\rho(x))\|_2 \leq \sigma_e^2 \}.$$

Although POCS has played an important role in nonparametric image reconstruction, it is not without limitations. Three of the most serious limitations are: First, it can handle only convex-type constraints, which prevents some effective but nonconvex-type constraints to be used. Secondly, POCS can be computationally expensive, because convex set projection operators are nonlinear in nature and sometimes involve constrained nonlinear optimization: for example, $P_{\Omega_5}$ for the convex set $\Omega_5$ defined above requires solving a quadratic programming step. Thirdly, POCS may converge very slowly and not necessarily to a unique solution when the projection operators are not contractive mapping; therefore, POCS can also converge a “bad” solution.
2.3.2 Parametric Methods

Parametric image models often take the following form:

\[ \rho(x) = \sum_n c_n \varphi_n(x), \]  

(2.9)

where \( \varphi_n(x) \) are the basis functions used to absorb any a priori information and \( c_n \) are the series coefficients chosen to match the measured data.

Selecting a set of “good” basis functions is essential for the model in Equation 2.9. A particular set of basis functions is given in the form of weighted complex sinusoids [5,6]:

\[ \varphi_n = C(x)e^{i2\pi n\Delta x}, \]  

(2.10)

where \( C(x) \) is a nonnegative function incorporating a priori information. With this set of basis functions, the model, known as the generalized series (GS) model [5,6], becomes

\[ \rho(x) = C(x)\sum_n c_n e^{i2\pi n\Delta x}. \]  

(2.11)

This model has several useful properties. Specifically, when no nontrivial a priori information is available, namely, \( C(x) = 1 \), Equation 2.11 automatically reduces to the conventional Fourier series model. This is desirable because the Fourier series model is indeed optimal in this case. On the other hand, if \( C(x) = \rho(x) \), the multiplicative Fourier series factor will be forced to unity by the data-consistency constraint, and a perfect reconstruction will result. In general, if \( C(x) \) is properly chosen, the new basis functions given in Equation 2.10 enable the GS model to converge faster than the Fourier series model. Therefore, within a certain error bound, fewer terms can be used to represent an image function than are required by the Fourier series method, leading to a reduction of the truncation artifact. The optimality of the GS model in Equation 2.11 can also be justified from the minimum cross entropy principle [7].

Selection of the weighting function \( C(x) \) is application dependent. For the limited data reconstruction problem, it was suggested [8] that \( C(x) \) be chosen to be a summation of boxcar functions as

\[ C(x) = \sum_{m=1}^M a_m \Pi \left[ \frac{x - \frac{1}{2}(\beta_m + \beta_{m+1})}{\beta_{m+1} - \beta_m} \right], \]  

(2.12)

where \( M \) represents the number of boxcar functions in the model, and \( \beta_m \) and \( a_m \) are the edge locations and amplitude of the \( m \)th boxcar, respectively. This function is particularly suitable for image functions containing sharp edges because they are explicitly built into the basis functions.
The weighting function \( C(x) \) can also be determined experimentally. A typical example is time-sequential imaging, which involves the acquisition of a time series of images, \( \rho_1(x), \rho_2(x), \ldots, \rho_L(x) \), from the same anatomical site. For many of this type of imaging experiments, the underlying high-resolution morphology in the desired image sequence does not change from one image to another. As a result, it is not necessary to acquire each of these images independently. Specifically, with the GS model, we first acquire one high-resolution (reference) data set with \( N \) encodings, followed by a sequence of reduced data set with \( M \) encodings. In the image reconstruction step, the high-resolution reference image \( \rho_{ref}(x) \) is used as the weighting function for the GS basis functions. That is, we set

\[
C(x) = |\rho_{ref}(x)| \quad (2.13)
\]

for the GS model when it is used for image reconstruction from the reduced data sets. After \( C(x) \) is known, the series coefficients \( c_n \) are determined by solving a set of linear equations from the data-consistency constraints. That is,

\[
D(n\Delta k) = \sum_{m=-N/2}^{N/2-1} c_m D^r[(n-m)\Delta k], \quad (2.14)
\]

where \( D(n\Delta k) = F\{C(x)\}(n\Delta k) \).

### 2.3.3 Application Examples

Constrained image reconstruction has been successfully used in several practical applications. This section discusses three specific examples: partial Fourier imaging, parallel imaging, and dynamic imaging.

#### Example 2.1: Partial Fourier Reconstruction

In partial Fourier imaging, \( k \)-space is sampled asymmetrically, say, \( D(n\Delta k) \) is measured for \( n \in N_{\text{data}} = \{-n_0, -n_0 + 1, \ldots, N-1\} \). Such a sampling scheme arises in MRI when a short echo time is used to avoid spin dephasing due to short \( T_2^* \) caused by local susceptibility changes or uncompensated motion effects. It is sometimes also used in the phase-encoding direction when an asymmetric set of phase-encoding measurements is acquired to reduce data acquisition time. Usually, \( n_0 \) is much smaller than \( N \), typically, \( n_0 = 16 \) or 32 with \( N \) being on the order of 128. The central \( k \)-space data are used first to obtain an phase estimate \( \hat{\phi}(x) \), which is then used as a constraint to get the final reconstruction. The phase-constrained reconstruction problem lends itself nicely to the POCS algorithm. Specifically, let

\[
\Omega_1 = \{ \rho(x) \mid \angle \rho(x) = \hat{\phi}(x) \} \quad (2.15)
\]

and

\[
\Omega_2 = \{ \rho(x) \mid F\{\rho(x)\} = D(n\Delta k), -n_0 \leq n \leq N-1 \}. \quad (2.16)
\]
Clearly, $\Omega_1$ contains all the images satisfying the predetermined phase constraint, whereas $\Omega_2$ contains all the images consistent with the measured data. The desired image $\rho(x)$ lies in the intersection of $\Omega_1$ and $\Omega_2$. That is, 

$$\rho(x) \in \Omega = \Omega_1 \cap \Omega_2,$$  

which can be found by alternating projections of an initial estimate onto these two sets. More specifically, 

$$\rho_{m+1}(x) = \varrho_1 \varrho_2 \{ \rho_m(x) \},$$  

where 

$$\varrho_1 \{ \rho(x) \} = | \rho(x) | e^{i\phi(x)}$$  

and 

$$\varrho_2 \{ \rho(x) \} = F^{-1} R F \{ \rho(x) \},$$  

in which $R$ is a data replacement operator defined as 

$$R \{ \hat{D}(n\Delta k) \} = \begin{cases} D(n\Delta k), & -n_0 \leq n \leq N - 1 \\ \hat{D}(n\Delta k), & \text{otherwise}. \end{cases}$$  

It is apparent that $\varrho_1$ projects any image function $\rho(x)$ onto $\Omega_1$, whereas $\varrho_2$ projects it onto $\Omega_2$. The initial condition $\rho_0(x)$ for Equation 2.18 is usually chosen to be the zero-filled Fourier reconstruction.

**Example 2.2: Data-Sharing Dynamic Imaging**

Constrained image reconstruction finds wide application in dynamic imaging. The keyhole and reduced encoding by generalized series reconstruction (RIGR) techniques [6,9,10] are two typical examples. A common feature of these two methods is that a high-resolution reference image and a sequence of reduced dynamic data sets (usually in central $k$-space) are collected. Assuming that $N$ encodings are collected for the reference data set and $M$ encodings for each of the dynamic data sets, a factor of improvement $N/M$ in temporal resolution (or imaging efficiency) is gained with this data acquisition scheme as compared to the conventional full-scan imaging method. In image reconstruction, the reference data is used to compensate for the loss of high-frequency data in the dynamic data sets. In keyhole, this is done in a straightforward fashion; that is, the unmeasured encodings of each dynamic data set are replaced directly by the corresponding reference data to create a “full-size” data set. A weakness of this data-sharing method is that any data
inconsistency between the dynamic and reference data sets will result in data truncation artifact and, as a result, dynamic image features are produced only at low resolution. With RIGR, image reconstruction is done using the GS model described in Section 2.2, in which the basis functions are determined by the reference data and the coefficients are determined by the dynamic data. This reconstruction algorithm can overcome the limited resolution problem with the keyhole method. It has been shown that with multiple references, RIGR can reconstruct dynamic features in a resolution close to that of the reference image [11].

2.4 REGULARIZED IMAGE RECONSTRUCTION IN PARALLEL MRI

2.4.1 BASIC RECONSTRUCTION METHODS

The Fourier image of the \(\ell\)th channel (ignoring the data truncation effects) is given by

\[
d_\ell(x) = \sum_{m=0}^{R-1} \rho(x - m\hat{W}) s_\ell(x - m\hat{W}),
\]

for \(\ell = 1, 2, \ldots, L\), and \(W/2 - \hat{W} < x < W/2\). Assuming that \(R \leq L\), we can solve for \(\rho(x)\) pixel by pixel from the earlier equations. More specifically, rewriting Equation 2.22 in matrix form

\[
S\tilde{\rho} = \tilde{d}
\]

where

\[
S = \begin{bmatrix}
s_1(x) & s_1(x - \hat{W}) & \cdots & s_1(x - (R-1)\hat{W}) \\
s_2(x) & s_2(x - \hat{W}) & \cdots & s_2(x - (R-1)\hat{W}) \\
\vdots & \vdots & \ddots & \vdots \\
s_L(x) & s_L(x - \hat{W}) & \cdots & s_L(x - (R-1)\hat{W})
\end{bmatrix}
\]

\[
\tilde{\rho} = \begin{bmatrix}
\rho(x) \\
\rho(x - \hat{W}) \\
\vdots \\
\rho(x - (R-1)\hat{W})
\end{bmatrix}, \quad \text{and} \quad \tilde{d} = \begin{bmatrix}
d_1(x) \\
d_2(x) \\
\vdots \\
d_L(x)
\end{bmatrix}.
\]

Equation 2.23 is known as the sensitivity encoding (SENSE) reconstruction formula [12], which can be derived from Papoulis’ generalized sampling theorem [13]. Clearly, perfect reconstruction of \(\rho(x)\) requires: (a) precise knowledge of \(s_\ell(x)\) to form, (b) to be nonsingular for \(W/2 - \hat{W} < x < W/2\), and (c) \(d_\ell(x)\) to be noiseless and not corrupted by the data truncation artifact.
In practice, Equation 2.23 is often solved in the least-squares (LS) sense or minimum-variance (MV) sense. The LS solution is given by

\[ \hat{\rho}_{LS} = (S^H S)^{-1} S^H \hat{d}, \]  

(2.24)

and the MV solution is given by [12]

\[ \hat{\rho}_{MV} = (S^H \Psi^{-1} S)^{-1} S^H \Psi^{-1} \hat{d}, \]  

(2.25)

where \( \Psi \) is the data noise covariance matrix. Some basic properties of the LS and MV solutions are summarized in the following remarks.

Remark 4: When \( s \) and \( \Psi \) are accurate, the variance of the reconstruction error due to data noise is given by

\[ \sigma_{LS}(x) = \sqrt{[(S^H S)^{-1} S^H \Psi S (S^H S)^{-1}]} \]  

for the LS solution, and

\[ \sigma_{MV}(x) = \sqrt{[S^H \Psi^{-1} S]} \]  

(2.26)

for the MV solution, where the subscript \( x \) denotes the index of the matrix corresponding to location \( x \).

Remark 5: The LS and the MV solutions are the same if the acceleration factor equals to the number of coils or noise is uncorrelated between coils, in which case there is no need to measure the noise covariance matrix.

Remark 6: The SNR of the MV solution is always greater or equal to that of the LS solution. The MV solution minimizes the variance of the reconstruction error vector \( E(\Delta \rho^H \Delta \hat{\rho}) = \text{trace}(\Delta \hat{\rho} \Delta \rho^H) \) over all possible estimators when the noise is Gaussian and over all linear unbiased estimators for non-Gaussian noise. Therefore, the mean-squared error of the MV solution is less than that of the LS solution.

The earlier results are based on the assumption that both \( S \) and \( \Psi \) are accurate. In practice, \( S \) and \( \Psi \) are estimated from experimental data, and any error in \( S \) (denoted as \( \Delta S \)) and/or in \( \Psi \) errors (denoted as \( \Delta \Psi \)) will contribute to \( \Delta \hat{\rho} \). Suppose that

\[ \| \Delta S \|_2 < \sigma_{\min}(S), \]

\[ \| \Delta \Psi \|_2 < \sigma_{\min}(\Psi), \]  

(2.28)
where \( \sigma_{\text{min}}(\cdot) \) denotes the minimum singular value of the matrix. It can be shown for the MV solution that \(^{14}\)

\[
\frac{\Delta \hat{\rho}_{\text{MV}}}{\hat{\rho}_{\text{MV}}} \leq \kappa(S_{\Psi}) \left( \frac{\|\Delta S_{\Psi}\|}{\|S_{\Psi}\|} + \frac{\|\Delta d_{\Psi}\|}{\|d_{\Psi}\|} \right),
\]

(2.29)

where the subscript \( \Psi \) denotes a matrix or vector pre-multiplied by \( \Psi^{-1/2} \), with \( \Psi^{-1/2} \) being defined as assuming that \( \Psi = V\Lambda^H \). The earlier result can be easily extended to the LS solution by setting \( \Psi = I \) in Equation 2.29.

### 2.4.2 Regularized Reconstruction Methods

The SENSE reconstruction (either \( \hat{\rho}_{\text{LS}} \) or \( \hat{\rho}_{\text{MV}} \)) is sensitive to \( \Delta \hat{d} \), \( \Delta S \), and \( \Delta \Psi \), especially when \( S \) is ill-conditioned. To desensitize the solution to data noise and model errors, regularization methods are often used. Tikhonov regularization is perhaps the most common regularization scheme, in which we form a weighted sum of the data misfit term \( \|S\hat{\rho} - \hat{d}\|^2 \) and a regularization term \( \|A(\hat{\rho} - \hat{\rho}_r)\|^2 \) using a weighting factor \( \lambda^2 \), and find the solution \( \hat{\rho}_{\text{reg}} \) that minimizes this sum, i.e.,

\[
\hat{\rho}_{\text{reg}} = \arg\min\{\|S\hat{\rho} - \hat{d}\|^2 + \lambda^2 \|A(\hat{\rho} - \hat{\rho}_r)\|^2\},
\]

(2.30)

where \( \lambda \) is often referred to as the regularization parameter and \( \hat{\rho}_r \) is a regularization image. A closed-form solution for \( \hat{\rho}_{\text{reg}} \) exists for \( L_2 \)-norm and is given by

\[
\hat{\rho}_{\text{reg}} = \hat{\rho}_r + (S^H S + \lambda^2 A^H A)^{-1} S^H (\hat{d} - S \hat{\rho}_r).
\]

(2.31)

Selecting “good” values for \( \lambda \) and \( \hat{\rho}_r \) is essential for this regularized reconstruction scheme. Although this is still a research problem, several algorithms have been proposed, which find useful practical applications. We will briefly review some of them to illustrate the concept.

### 2.4.2.1 Construction of \( \hat{\rho}_r \)

There are basically three schemes to construct \( \hat{\rho}_r \): (a) setting \( \hat{\rho}_r = 0 \), (b) recycling an initial SENSE reconstruction to create \( \hat{\rho}_r \), and (c) collecting additional data to generate \( \hat{\rho}_r \).

Scheme (a) corresponds to, perhaps, the simplest version of the Tikhonov regularization scheme. It was used in Reference 15 with some success. In scheme (b), the conventional SENSE algorithm is used to obtain an initial reconstruction, which is then filtered by a median filter to suppress any residual aliasing artifacts \(^{16}\). However, if the matrix \( S \) is highly ill conditioned within a large region, the filtering step may not be effective in suppressing the aliasing artifacts. Scheme (c) acquires additional \( k \)-space center lines at the Nyquist frequency.
rate, known as the autocalibration scan [17,18,19] and uses these data to reconstruct low-resolution regularization image [20]. A high-resolution regularization image can also be created from these data using the GS model. Details of the algorithm can be found in Reference 21.

Figure 2.2 shows a set of regularized reconstructions with different regularization images from real experimental data acquired with four receiver coils and $R = 4$. As can be seen, different regularization images can affect the final reconstruction.

2.4.2.2 Selection of $\lambda$

A straightforward way to select the regularization parameter is to set $\lambda$ heuristically as a constant over the entire image. This method is not effective because the condition of $S$ varies at different locations. A more elaborate way is to select $\lambda$ adaptively using traditional regularization methods such as the $L$-curve or the generalized cross-validation (GCV) methods [22]. The $L$-curve method was used in parallel imaging with some success [20]. The GCV method works well in general but sometimes gives biased results if the noise $\Delta d$ is highly correlated [22].
A significant weakness of both the conventional L-curve and the GCV methods lies in the fact that they choose $\lambda(x)$ independently for different spatial locations. This problem was addressed in Reference 21 with an algorithm to select $\lambda(x)$ jointly. Specifically, the algorithm first sets $\lambda(x)$ to be within $[\lambda_{\min}, \lambda_{\max}]$, and then forms $\lambda(x)$ as a linear function of the local condition number of $S$, i.e.,

$$\lambda(x) = \alpha \kappa(S) + \beta,$$

for $W/2 - \hat{W} < x < W/2$. This scheme is based on the consideration that the larger the $\kappa(S)$, the heavier the regularization is needed for Equation 2.31. To determine $\alpha$ and $\beta$, Equation 2.32 is rewritten as

$$\lambda(x) = \frac{\kappa(S) - \kappa_{\min}}{\kappa_{\max} - \kappa_{\min}} (\lambda_{\max} - \lambda_{\min}) + \lambda_{\min},$$

where $\kappa_{\max}$ and $\kappa_{\min}$ are the maximum and the minimum condition numbers of all $S$, and $\lambda_{\min}$ and $\lambda_{\max}$ are determined by

$$\lambda_{\min} = \arg \min \left\{ \frac{\max \sigma_i}{\min (\sigma_i + \lambda^2 / \sigma_i)} < K \right\},$$

and

$$\lambda_{\max} = \arg \max \left\{ \sum_x \| S \hat{\rho}_{\text{reg}}(\lambda) - \hat{d} \| \leq \epsilon \right\},$$

where $\sigma_i$ is the $i$th singular value of $S$, and $K$ and $\epsilon$ are user-specified constants. Details of the algorithm can be found in [21].

Figure 2.3 shows a set of exemplary regularized reconstructions with different regularization parameters from real experimental data acquired with four receiver coils and an acceleration factor of four. The importance of regularization parameter can be appreciated by comparing the results in (a) to (d).

### 2.4.2.3 Sensitivity Analysis

An upper bound for the sensitivity of the regularized solution to data noise and model error is given by

$$\frac{\| \Delta \hat{\rho}_{\text{reg}} \|}{\| \hat{\rho}_{\text{reg}} \|} \leq \sigma_{\max} \left( \frac{\| \Delta S \|}{\| S \|} + \frac{\| \Delta \hat{d} \|}{\| d \|} \right) + \frac{\lambda^2}{\sigma_i^2 + \lambda^2} \frac{\| \hat{\rho}_{\text{reg}} - \hat{\rho} \|}{\| \hat{\rho} \|},$$

where $\sigma_{\max}$ denotes the largest singular value of matrix $S$ and

$$q = \arg \min \left\{ \frac{\sigma_j + \lambda^2}{\sigma_j} \right\}.$$
As expected, the result in Equation 2.29 is a special case of Equation 2.36 when $\lambda = 0$. Comparing Equation 2.36 and Equation 2.29 yields

$$k(S_\lambda) = \frac{\sigma_{\max}}{\sigma_q + \frac{\lambda}{\sigma_q}},$$

which can be regarded as the effective condition number of the regularized reconstruction. Clearly, $k(S_\lambda)$ is reduced by increasing $\lambda$.

### 2.4.3 Application Example

An example is shown in Figure 2.4, where the data were collected using three coils in a dynamic contrast-enhanced MRI experiment. In addition to the usual SENSE data, eight encodings were collected at the Nyquist rate in
central $k$-space for each data frame, from which $\tilde{\rho}_c$ was derived using the GS model. As can be seen, the SENSE reconstruction using the proposed algorithm (Figure 2.4b) is significantly better than that from the standard SENSE algorithm (Figure 2.4a).

### 2.5 CONCLUSION

Image reconstruction from limited Fourier data is a classical problem in tomographic imaging. Although a general solution to this problem is not available, a number of practical techniques have emerged, which can provide optimal (or close-to-optimal) solutions to a particular application problem, leading to significant improvements in image quality. This chapter provided a tutorial discussion of some representative techniques, including parametric and nonparametric methods for superresolution image reconstruction from limited Fourier data and regularization methods for image reconstruction from multichannel undersampled Fourier data. The chapter is also intended to provide some basic background knowledge of the area for the reader to apply these techniques to particular problems or to further improve them.
REFERENCES


