Reviews of Magnetic Resonance in Medicine, Vol. 4, pp. 67–185, 1992
Printed in the USA.
0883-8291/92 $5.00 + .00

Constrained Reconstruction Methods in MR Imaging

Zhi-Pei Liang,* Fernando E. Boada,† R. Todd Constable,‡ E. Mark Haacke,§ Paul C. Lauterbur,* and Michael R. Smith¶

*University of Illinois at Urbana-Champaign
†MGH-Harvard University
‡Yale University
§Case Western Reserve University
¶University of Calgary

OUTLINE

Address correspondence to Z.-P. Liang, Biomedical Magnetic Resonance Laboratory, University of Illinois at Urbana-Champaign, 1307 West Park Street, Urbana, IL 61801.

INTRODUCTION

Aims of the Review

In magnetic resonance imaging (MRI), spatially-resolved information, such as spin density, relaxation parameters, and chemical shift information, can be encoded in the measured data in a variety of ways. After more than a decade of evolution in data acquisition strategies (pulse sequences), the Fourier-encoding scheme, in which the spatial origins of this information are encoded as frequency and phase differences in the NMR signals, has become most popular. The purpose of image reconstruction is to decode the measured signals to identify various spatial frequency components present and map them to the corresponding spatial locations. This process, often called spectral analysis in signal processing, is usually accomplished through the use of the fast Fourier transform (FFT) algorithm. This method is computationally efficient and produces reasonable results when the Fourier space (k-space) is sampled sufficiently. However, there are three well-known inherent limitations that make it less desirable for many MRI applications. The most serious one is due to the convolution effect from finite sampling or data truncation. This effect manifests itself as blurring and ringing in the image domain, obscuring and distorting the reconstructed images. The second important limitation is its inability to resolve fine image details (pixel-limited resolution). With the Fourier method, spatial resolution is limited to roughly the reciprocal of the frequency interval over which the sampled (measured) data are available. When the sampling window is small (or equivalently, the number of phase-encoding or frequency-encoding steps is small), the resultant large pixel size will give rise to significant partial-volume artifacts. The third limitation lies in its signal-to-noise inefficiency, namely, requirements for improving image signal-to-noise ratio (SNR) and spatial resolution are mutually exclusive and, therefore, an improvement in image resolution with extended k-space sampling will lead to a concomitant loss of image SNR.

The first two limitations of the Fourier reconstruction method are particularly troublesome when the data sets are small, as is frequently the case in practice. For example, in 3-D imaging, 64 or fewer phase-encoding steps are often used in the third dimension to

reduce data acquisition time. Even with the standard 2-D imaging, the number of phase-encoding steps is often reduced to 128 or fewer in many imaging applications because of imaging time or SNR considerations. In spectroscopic imaging, where only 8 or 16 phase-encoding steps are usually acquired in each spatial direction, this limited-data problem is much more serious. Direct applications of the Fourier method yield results that contain serious artifacts, preventing reliable interpretation of the results.

To alleviate one or more of these problems with the Fourier approach, many alternative reconstruction methods have been proposed in various imaging areas. In particular, the class of constrained methods emerging within the last decade appears very promising. Although all these methods are reported to be somewhat successful, neither their possibilities nor their pitfalls are well understood. The main purpose of this review is to give a systematic discussion and evaluation of these methods with respect to their applications in MRI. We attempt not only to summarize the available material scattered in the literature, but also to give a reasonably complete discussion of the theory and algorithms for each method. An emphasis will also be placed upon evaluation and comparison of these methods so that their relative strengths and weaknesses can be fairly assessed. As a whole, this review may provide an introduction to constrained image reconstruction for an MRI scientist who needs it, but does not know how to get started. It should also serve as a good beginning for the scientist or student preparing for research aimed at improved image reconstruction methods. It contains a fairly complete list of references that will enable the researcher to quickly identify many of the important contributions.[1]

Finite-Fourier Imaging

Since the introduction of the gradient-encoding concept by Lauterbur in the early 70's (128), frequency-encoding and phase-encoding of the spatial origin of an MR signal in an object have become standard in multi-dimensional MR imaging and spectroscopy. This paper mainly concerns the corresponding image reconstruction problem. For simplicity, we consider here only the one-dimensional (1-D) reconstruction problem, assuming that reconstruction along the other directions has been done or is to be done. This assumption is valid since higher-dimensional reconstruction is usually accomplished by cascaded 1-D processing, although some 2-D methods are emerging, such as those discussed in (71,218). In spectroscopic imaging, data truncation is used in all spatial directions and the discussion there will be naturally cast in the higher-dimensional settings.

In k-space notation (36,238), the MR signal $\hat{s}(k)$ collected along either the frequency-encoding or phase-encoding direction can be described by the following Fourier integral

$$\hat{s}(k) = \int_{-\infty}^{+\infty} \rho(x) e^{-i2\pi kx} dx + \varepsilon(k) = s(k) + \varepsilon(k), \tag{1}$$

where $\rho(x)$ is the desired image function (spin density function weighted by relaxation time effects), and $\varepsilon(k)$ represents the measurement errors from various sources, including

[1]Many of the reconstruction algorithms discussed in this review are being incorporated into a general purpose image processing package developed at the National Center for Supercomputing Applications (NCSA) and the Biomedical Magnetic Resonance Laboratory (BMRL) of University of Illinois; interested readers should contact Z.-P. Liang or P. C. Lauterbur for further information.

random noise. In practice, the image function, $\rho(x)$, is a complex function and has energy only in a finite spatial region $\mathcal{B} = (-L_x/2, L_x/2)$, namely, $|\rho(x)| = 0$ for $x \notin \mathcal{B}$, where L_x is the length of the excited region of an object being imaged and $\mathcal{B}$ corresponds to part or all of the field-of-view (FOV) of the image. Such a function is often called *spatial-support-limited* or simply *support-limited* in signal processing. The measured signal $\hat{s}(k)$ is assumed to be uniformly sampled at $k = n\Delta k$, for $n \in \mathcal{N}_{data}$, with the sampling interval Δk satisfying the Nyquist criterion ($\Delta k \leq 1/L_x$). When $\mathcal{N}_{data}$ is a finite integer set, we have the *finite-Fourier imaging problem,* so called because a finite number of sampled data points is collected in Fourier space. In this review, we will consider two special cases of finite-Fourier imaging: *symmetrically truncated Fourier imaging* and *asymmetrically truncated Fourier imaging*. In the first case, the Fourier space is symmetrically sampled such that $\mathcal{N}_{data} = \{-N, -N + 1, \cdots, N - 1\}$ for a small N, (64 or less, for example). In the second case, the sampling is asymmetric, such that $\mathcal{N}_{data} = \{-n_0, -n_0 + 1, \cdots, N - 1\}$ for $n_0 \ll N$ (or $n_0 \gg N$, but for simplicity, we assume $n_0 \ll N$ throughout this review). As illustrated in Fig. 1, the first imaging problem is characterized by the loss of high spatial frequency data and will be called *truncated-Fourier imaging;* the second imaging problem is characterized by the sampling asymmetry and is often called in the literature *half-Fourier, one-sided* or *partial-Fourier* imaging. In order to avoid unnecessary proliferation of terminologies leading to possible confusion, we will attempt to make consistent use of the term partial-Fourier imaging for this data acquisition scheme. For convenience, we will also introduce N_{data} to denote the total number of measured data points, which is equal to $2N$ in the symmetric case and to $(n_0 + N)$ in the asymmetric case.

Finite sampling (symmetric or asymmetric) is common in many imaging problems. In MRI, often only a limited set of phase-encoded measurements is taken due to time and physical (mainly gradient strength) constraints. This naturally leads to the finite-Fourier imaging problem. It also occurs in the read-out (or frequency-encoding) direction when few data points are acquired for each echo, as is often the case with echo-planar imaging or snapshot imaging. Partial-Fourier sampling is often associated with short-echo-time imaging or flow imaging, where motion-compensated asymmetric gradient echos are acquired to reduce spin dephasing and motion artifacts.

In general, an infinite number of sampled data points of $\hat{s}(k)$ is necessary to reconstruct $\hat{\rho}(x)$ exactly (in the zero mean-squared error sense). When a finite number of sampled data points is available, the solution of integral equation (1) is not unique, i.e., there will be

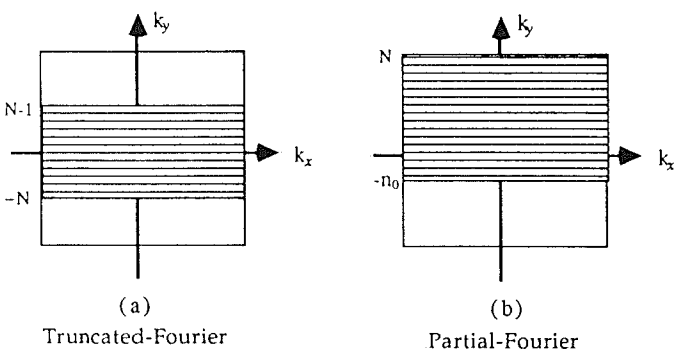

Figure 1. *k*-space sampling of finite-Fourier imaging.

infinitely many reconstructions of $\hat{\rho}(x)$ consistent with the measured data values. Some of these reconstructions will be reasonably good, while others will not; the data constraints by themselves will not automatically lead to a good reconstruction unless the number of data points is large. The goal of image reconstruction for finite-Fourier imaging is, therefore, not simply to find an image function satisfying Eq. (1), but instead to find an optimal one with the best possible spatial resolution and signal-to-noise. The constrained reconstruction methods, to be reviewed in this paper, represent the current approaches to achieving this objective.

The Concept of Constrained Reconstruction

For years, people were led to believe that information beyond the measurement cutoff frequency were not recoverable, e.g., the Rayleigh resolution limit for diffraction-limited imaging systems (15,19,23,25,60,90,91). Although the information is not apparent in the measured data, people have now learned how to look elsewhere for the additional information required to restore those frequencies. 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. The fundamental difference between the conventional Fourier reconstruction method and modern constrained reconstruction methods is, therefore, that the former treats the unmeasured data as zeros and the latter recovers them based on *a priori* information. Although constrained data processing methods have been extensively used for decades in various fields, application of the constrained reconstruction concept to MRI is very recent. The first successful effort was perhaps due to Smith et al. (214) 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 worth the effort. In particular, the ability to reduce the Gibbs artifact and improve image resolution is very desirable and can produce results unmatched by the traditional unconstrained Fourier methods. It can be anticipated that with further improvements in performance and better "public" understanding of their strengths, some of these methods will find their way into general practical applications for improving the efficiency and reliability of MR experiments.

In this review, we will make extensive use of the word "constraint," meaning *a priori* information, bounds, or parametric models placed upon the reconstruction. Constraints are available in practice in a variety of forms. For example, one can use the positivity constraint if the image function is known to have only nonnegative values, or one can force the spatial bound to the reconstruction if the energy distribution of the object being imaged is known to vanish outside a well-defined spatial region. Sometimes, even the data acquisition can be intelligently designed so that one set of the data can be used to constrain the others, such as the case with constrained spectroscopic imaging, to be discussed later. In principle, any valid prior knowledge about the object being imaged can be used to constrain the reconstruction for improved accuracy. In any case, the best reconstruction should result from incorporating the maximum amount of *valid* prior and experimental information available. However, it should also be emphasized that not all constraints are useful for a particular problem. If not used properly, *a priori* constraints may bias the interpretation of the limited amount of experimental data, which, at best,

over-constrains the problem, preventing discovery of new image features (or important symptoms) — a departure from the fundamental goal of an imaging experiment, and at worst gives rise to new image artifacts, possibly leading to false diagnoses (or medical decisions). One of the important tasks of this paper is to reveal whether any of the proposed methods fall into this category.

Depending on what the constraints are and how they are used in a reconstruction method, we roughly classify the existing methods into three major groups: *phase-constrained methods, nonparametric methods,* and *parametric modeling methods*. This classification, to some extent, is arbitrary, but will help to emphasize the key features of individual methods. For example, the phase-constrained methods, by definition, force the final reconstruction with a certain predetermined phase, and can be classified as nonparametric methods or parametric methods depending on what types of phase constraints are used. But we will put these methods in a separate category and review them in connection with image reconstruction from partial-Fourier data. In the group of nonparametric methods, we basically include all the methods that do not involve parametric modeling of the desired image function. We distinguish parametric methods from nonparametric methods, since the underlying constraints are considerably different. Nonparametric constraints permit the use of the conventional Fourier series model for an 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 artifacts. Parametric modeling methods, on the other hand, represent the image function in terms of a set of parametrized basis functions, rather than the nonparametrized 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. We will discuss this type of model-based extrapolation reconstruction in the parametric modeling section, mainly for the sake of emphasizing the nature of the constraints, rather than maintaining the accuracy of classifications of the reconstruction procedures.

It is also worthwhile to point out that in addition to changing the way that an image is synthesized (reconstructed) from the measured data, *a priori* constraints can also dictate the data acquisition process to further optimize the achievable spatial resolution and SNR efficiency. A simple example is the use of non-uniform sampling in k-space, such as the case with the SLIM (Spectral Localization by IMaging) method to be discussed later. In this review, we will limit ourselves to the discussion of only the reconstruction aspects of utilizing *a priori* constraints.

Organization of the Review

In reviewing an image reconstruction method, this paper mostly follows a characteristic format: First, the basic theory and concept are presented, and then the computational algorithm is described, followed by a section discussing and summarizing the important properties of the methods, with some reconstruction examples mainly from a phantom data set. The whole review has six major sections. The *Fourier Reconstruction Method* is devoted to background material, briefly reviewing some important properties of the con-

ventional Fourier reconstruction method, including Gibbs's artifact and SNR efficiency. In the *Phase-Constrained Reconstruction* section, a specialized class of constrained methods for image reconstruction from partial-Fourier data is reviewed with applications to short echo time imaging. In the next two sections, methods for reducing the Gibbs artifact and improving resolution in finite-Fourier imaging are discussed. The nonparametric methods that make use of constraints, such as finite spatial support, maximum entropy, linear predictability, and image smoothness are reviewed in the *Nonparametric Constrained Methods,* while parametric modeling methods, including ARMA (autoregressive moving average), LPA (localized polynomial approximation), and GS (generalized series), are covered in the *Parametric Modeling Methods* section. Since constrained spectroscopic imaging is still a relatively young area and only a few methods have been proposed, the *Constrained Spectroscopic Imaging* section will focus on reviewing the problems, difficulties and some existing solutions, so as to stimulate more research in this extremely important area. In the last section, various methods will be further discussed and compared based on several practical data sets, followed by the general conclusion of the whole review.

FOURIER RECONSTRUCTION METHOD

Introduction

The commonly used Fourier reconstruction method is based on the well-established Fourier inversion theory. In general, values of $s(k)$ need to be known for all the spatial frequencies k to reconstruct exactly the image function $\rho(x)$ from the continuous inverse Fourier transform

$$\rho(x) = \int_{-\infty}^{\infty} s(k)e^{i2\pi kx}dk. \tag{2}$$

For a support-limited image function, this requirement can be considerably relaxed, based on the Fourier series expansion theory in which the periodic extension[2] of $\rho(x)$ can be expressed as

$$\tilde{\rho}(x) = \Delta k \sum_{n=-\infty}^{\infty} s(n\Delta k)e^{i2\pi n\Delta kx}. \tag{3}$$

It is known that for $x \in \text{FOV} = [-1/(2\Delta k), 1/(2\Delta k)]$, $\tilde{\rho}(x)$ converges to $\rho(x)$ when $\rho(x)$ is a continuous function, or to $[\rho(x^+) + \rho(x^-)]/2$ when $\rho(x)$ is piecewise continuous with only step discontinuities (31,33,176). For a given finite set of measured data $\{s(n\Delta k), n$

[2]Periodic extension of a $[-L_x/2, L_x/2]$-support-limited function $\rho(x)$ with a period $\tilde{L}_x = 1/\Delta k \geq L_x$ is defined as

$$\begin{cases} \tilde{\rho}(x) = \rho(x), & x \in [-L_x/2, L_x/2] \\ \tilde{\rho}(x) = 0, & -\tilde{L}_x/2 \leq x < -L_x/2 \text{ or } L_x/2 < x \leq \tilde{L}_x/2 \\ \tilde{\rho}(x) = \tilde{\rho}(x + n\tilde{L}_x), & \forall n. \end{cases}$$

$\in \mathcal{N}_{data}\}$, one can then obtain an approximated image function from the corresponding truncated Fourier series as

$$\rho_{DFT}(x) = \Delta k \sum_{n \in \mathcal{N}_{data}} s(n\Delta k) e^{i2\pi n \Delta k x}, \tag{4}$$

in which the unavailable high spatial frequency components, i.e., $s(n\Delta k)$, for $n \notin \mathcal{N}_{data}$, are treated as being zero. This image function is what people commonly get using a discrete Fourier transform (DFT) reconstruction procedure, and will be simply called Fourier reconstruction in distinction to reconstructions from other methods to be discussed later.

In practice, $\rho_{DFT}(x)$ is evaluated at $x = m\Delta x$, $m \in \mathcal{N}_{imag} = \{-N_{imag}/2, \cdots, N_{imag}/2 - 1\}$, inside the FOV for display. If the digital pixel size Δx equals $1/(N_{data}\Delta k)$, this set of image values can be calculated efficiently by applying an FFT algorithm to the data. The reconstruction thus obtained is often called FFT reconstruction.[3] If a smaller digital pixel size is desired, for example, $\Delta x = 1/(N_{imag}\Delta k)$ for $N_{imag} > N_{data}$, $s(n\Delta k)$ is usually zero-filled to N_{imag} points in the unsampled region before the FFT is applied. The result is thus called zero-padded FFT reconstruction. Zero-filling can be done in a variety of ways with the same effect on the magnitude reconstruction. For example, the zeros can be padded at both ends as $\{0, \cdots, 0, \{s(n)\}, 0, \cdots, 0\}$ or at one end as $\{0, \cdots, 0, \{s(n)\}\}$ or $\{\{s(n)\}, 0, \cdots, 0\}$ and the only difference in the resultant reconstruction will be in the phase images (33). Zero-padded reconstruction may sometimes look better because of the interpolation effects, but no resolution is gained by zero-filling. When padding with a large number of zeros, specialized FFT algorithms, such as those in (167,217), can be used for improved computational efficiency.

In the rest of this section, several important properties of the Fourier reconstruction method are reviewed. This will provide the necessary background for understanding why and how the constrained methods, to be discussed later in this review, are developed. For simplicity, we consider only the symmetric case of finite sampling, i.e., $\mathcal{N}_{data} = \{-N, -N + 1, \cdots, N - 1\}$, and $N_{data} = 2N$, but with little modification, most of the conclusions apply to the asymmetric case as well.

Gibbs' Artifact

In this section, we examine and characterize the effects of data truncation on the Fourier reconstruction. Making use of the well-known convolution theorem (31,176), it is easy to derive from Eqs. (3) and (4) that

$$\rho_{DFT}(x) = \int_{-L_x/2}^{L_x/2} \tilde{\rho}(\tau) h(x - \tau) d\tau = \int_{-\infty}^{+\infty} \rho(\tau) h(x - \tau) d\tau$$

or

$$\rho_{DFT}(x) = \rho(x) \bigstar h(x), \tag{5}$$

[3]The terms Fourier reconstruction and FFT reconstruction will be used interchangeably in this review.

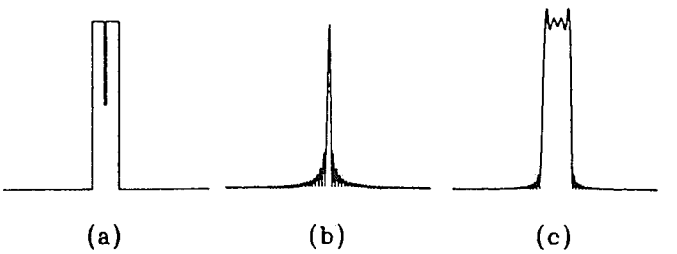

Figure 2. Illustration of the truncation artifact: (a) The true image function, $\rho(x)$; (b) convolution kernel, $h(x)$; and (c) The Fourier reconstruction, $\rho_{DFT}(x)$.

where the convolution kernel is given by[4]

$$h(x) = \Delta k \sum_{n=-N}^{N-1} e^{i2\pi n\Delta kx} = \Delta k \frac{\sin(2\pi N\Delta kx)}{\sin(\pi\Delta kx)} e^{-i\pi\Delta kx}. \tag{6}$$

Therefore, approximating a support-limited function $\rho(x)$ by the discrete Fourier transform $\rho_{DFT}(x)$ is equivalent to convolving the original function, $\rho(x)$, with a kernel function, $h(x)$, called the *point spread function* (PSF). Due to the oscillatory nature of $h(x)$, Fig. 2b, this convolution merges closely-spaced features together and gives rise to spurious ringing in $\rho_{DFT}(x)$, as illustrated in Fig. 2c.

The convolution effect manifests itself differently for functions with different continuity properties. When $\rho(x)$ is a smooth function, $\rho_{DFT}(x)$ uniformly converges to $\rho(x)$ as $N \rightarrow \infty$ for $x \in$ FOV. More precisely, if $\rho(x)$ has continuous derivatives up to and including the p^{th}-order, then $\rho_{DFT}(x)$ approaches $\rho(x)$ on the order of $1/N^{p+1}$, namely, as N goes to infinity $\|\rho_{DFT}(x) - \rho(x)\|$ approaches zero as fast as $1/N^{p+1}$ does. Therefore, for smooth functions, the relative amount of truncation induced by finitely sampling the data will be small and, as a consequence, the amount of spurious ringing due to convolution with $h(x)$ will be relatively insignificant (31,176).

On the other hand, if $\rho(x)$ contains step discontinuities, the convergence behavior in the neighborhood of a point of discontinuity will be "anomalous." Assuming that $\rho(x)$ has a step discontinuity at $x = x_0$ such that $\rho(x_0^+) > \rho(x_0^-)$, the maximum overshoot location x_N^{os} in $\rho_{DFT}(x)$ approaches x_0^+ as $N \rightarrow \infty$, and

$$\lim_{N\rightarrow\infty} [\rho_{DFT}(x_N^{os}) - \rho(x_0^+)] \approx 9\% [\rho(x_0^+) - \rho(x_0^-)]. \tag{7}$$

---

[4]Note that in the continuous case, the convolution kernel is given by

$$h(x) = \frac{\sin(\pi W_k x)}{\pi x} = W_k \text{sinc}(\pi W_k x)$$

where W_k is often called the aperture size in optical imaging systems (25). Its effect is similar to that of the discrete counterpart.

Similarly, for the maximum undershoot, we have

$$\lim_{N\to\infty} [\rho_{DFT}(x_N^{us}) - \rho(x_0^-)] \approx -9\% \, [\rho(x_0^+) - \rho(x_0^-)]. \tag{8}$$

Therefore, as $N \to \infty$, the value of the maximum overshoot (undershoot) does not tend to zero, but instead tends to a finite limit. The existence of this finite, nonzero, limiting value of the overshoot (undershoot) is due to the nonuniform convergence of $\rho_{DFT}(x)$ to $\rho(x)$ in the vicinity of discontinuous points of $\rho(x)$. This nonuniform behavior of the limit $\rho_{DFT}(x) \to \rho(x)$ as $N \to \infty$ for $x \in$ FOV is called the *Gibbs' phenomenon*.

In some MR literature, the terms truncation artifacts and Gibbs phenomenon are often used exchangeably. However, there exists a fundamental difference between these two concepts. The former refers to the error (resolution limitation and spurious ringing) resulting from data truncation, and the latter characterizes only the limiting behavior of this error as $N \to \infty$. In this review, we attempt to distinguish these two, but will use the term Gibbs ringing (artifact) to refer to the spurious oscillations resulting from data truncation.

In summary, with the Fourier reconstruction method, data truncation causes Gibbs ringing. This ringing has the following characteristics.

1. The maximum undershoot or overshoot is about 9% of the intensity discontinuity and is independent of the number of data points used.
2. The frequency of the spurious ringing increases with more data points being used (The widths of the main-lobe and side-lobes of $h(x)$ are equal to $2/(N_{data}\Delta k)$ and $1/(N_{data}\Delta k)$, respectively). Therefore, the distance over which the ringing is significant is reduced as resolution is increased. This is why in practice, when a large number of data points is used, the spurious ringing does not cover an appreciable distance in the Fourier reconstruction. Note that when zero-padding is used, the ringing frequency is a function of the original number of data points N_{data} not including the padded zeros. Zero-padding frequently makes the ringing, actually present, more visually apparent, but does not introduce the ringing.

Data Windowing Methods

In order to remove the spurious ringing with the simple Fourier reconstruction method, many filtering (windowing) schemes were proposed in the sixties and the early seventies (89). Some new window functions are also discussed in the recent literature (2). These methods modify the shape of the point spread function $h(x)$ to reduce the amplitudes of its side-lobes by properly weighting the available data, but always at the expense of widening its main-lobe, thus reducing resolution. In general, the windowed Fourier reconstruction is expressed as

$$\rho_{WFT}(x) = \Delta k \sum_{n \in \mathcal{N}_{data}} s(n\Delta k) w(n) e^{i2\pi n \Delta k x}, \tag{9}$$

where $w(n)$ is a pre-designed window (filter) function. For example, in the symmetric sampling case, the popular Hamming window function is defined as follows:

$$w(n) = 0.54 + 0.46 \cos(n\pi/N), \qquad -N \leq n < N. \tag{10}$$

In the asymmetric case, the data is either filtered on one side by

$$w(n) = 0.54 + 0.46 \cos(n\pi/n_0), \qquad -n_0 \leq n \leq 0, \tag{11}$$

or on both sides by (171)

$$w(n) = 0.54 + 0.46 \cos[(n - n_0)\pi/(2n_0)], \qquad -n_0 \leq n \leq n_0. \tag{12}$$

The Hamming window gives approximately the minimum sidelobe levels. Thus, the magnitude of the spurious ringing with the Hamming-filtered Fourier reconstruction is about minimal.

Another useful window function is defined through the so-called discrete prolate sequence (174), which is an eigenvector of the following linear system

$$\sum_{m=-N}^{N} \frac{\sin(n - m)\pi\alpha}{(n - m)\alpha} w(m) = \lambda w(n), \qquad |n| \leq N. \tag{13}$$

The eigenvector corresponding to the maximum eigenvalue λ_{max} defines a window function with the maximum energy concentration in a specified interval $(-\alpha, \alpha)$. This filter, therefore, provides an effective means for reducing ringing with a minimal loss of image resolution if the filter parameter α is chosen properly. Application of this filter to MR image processing has been recently explored by Weaver et al. (244).

For a thorough review of many of the filter functions and their properties for spectral analysis, one is referred to (89).

Resolution Limitation

Spatial resolution is an imaging parameter of considerable interest in virtually all imaging systems. Qualitatively, spatial resolution is defined as the smallest separation, δx, of two point-sources necessary for them to remain resolvable in the image. Equivalently, spatial resolution of an imaging system characterizes how accurately the dimensions, typically the edge locations, of an object can be determined from the image. In a Fourier imaging system, δx depends on the effective width of the point spread function $h(x)$, which is usually taken to be the duration of the equivalent rectangular pulse approximating $h(x)$ in a cycle, namely,

$$\delta x = \frac{1}{h(0)} \int_{-L_x/2}^{L_x/2} h(x)dx = \frac{1}{N_{data}\Delta k}. \tag{14}$$

Sometimes, δx is defined as the full-width at half maximum of $h(x)$. If the data $s(k)$ are filtered with a window function $w(k)$ to reduce the Gibbs ringing, as discussed in the previous section, then

$$\delta x = \frac{1}{\sum_{n\in\mathcal{N}_{data}} w(n\Delta k)} \int_{-L_x/2}^{L_x/2} \sum_{n\in\mathcal{N}_{data}} w(n\Delta k) e^{i2\pi n\Delta k x} dx$$

$$= \frac{1}{\sum_{n\in\mathcal{N}_{data}} [w(n\Delta k)/w(0)]\Delta k} \geqslant \frac{1}{N_{data}\Delta k} . \tag{15}$$

This equation indicates that the resolution has decreased because of the filtering operation. Therefore, for any reconstruction algorithms that do not lead to a recovery of some of the missing high spatial-frequency data, the best spatial resolution that can be obtained is

$$\delta x_{opt} = \frac{1}{N_{data}\Delta k} . \tag{16}$$

In other words, *a priori* information is necessary to overcome this resolution limitation to achieve what is often called *superresolution reconstruction.*

The right-hand side of Eq. (16) is usually called the Fourier pixel size Δx_f in contrast to the usual image digital pixel size Δx_d or simply Δx. Note that Δx can be made arbitrarily small by using zero-padding or other interpolation schemes, but the image resolution is fundamentally limited to the Fourier pixel size. Therefore, in order for two point sources to be distinguishable, their separation has to be larger than Δx_f. Another implication of Eq. (16) is that, if we are to determine the object edge locations from the image thus obtained, an accuracy of Δx_f is the best we can obtain. Eq. (16) also indicates that δx and N_{data} cannot be reduced simultaneously, namely, we cannot improve image resolution and reduce the number of measurements at the same time. This assertion is known as the *uncertainty principle* of the Fourier transform.

Signal-to-Noise Characteristics

Measurement noise is also an important limiting factor for obtaining high spatial resolution in practical imaging systems. If the noise term, $\varepsilon(k)$, in Eq. (1) is not zero, the noise in the image domain will be $e(m\Delta x)$, which is given by

$$e(m\Delta x) = \Delta k \sum_{n=-N}^{N-1} \varepsilon(n\Delta k) e^{i2\pi mn\Delta k\Delta x} \tag{17}$$

and

$$\hat{\rho}_{\mathrm{DFT}}(m\Delta x) = \rho_{\mathrm{DFT}}(m\Delta x) + e(m\Delta x). \tag{18}$$

Furthermore, if we assume that $\varepsilon(k)$ is an ergodic, stationary, uncorrelated, white noise process with zero mean and standard deviation δ_{data}, then $e(m\Delta x)$ has the following properties:

1. It has a zero mean, namely,

$$E\{e(m\Delta x)\} = E\left\{\Delta k \sum_{n\in\mathcal{N}_{data}} \varepsilon(n\Delta k)e^{i2\pi nm\Delta k\Delta x}\right\} = 0, \tag{19}$$

2. Its variance is

$$\begin{aligned}\delta^2_{imag} &= E\{[e(m\Delta x) - E\{e(m\Delta x)\}][e(m\Delta x) - E\{e(m\Delta x)\}]^*\} \\ &= \Delta k^2 \sum_{p\in\mathcal{N}_{data}} \sum_{q\in\mathcal{N}_{data}} E\{\varepsilon(p\Delta k)\varepsilon^*(q\Delta k)\}e^{i2\pi(p-q)m\Delta k\Delta x} \\ &= \Delta k^2 \sum_{p\in\mathcal{N}_{data}} E\{\varepsilon(p\Delta k)\varepsilon^*(p\Delta k)\} = N_{data}\Delta k^2\sigma^2_{data},\end{aligned} \tag{20}$$

3. It is uncorrelated, i.e., $E\{e(m\Delta x)e^*(n\Delta x)\} = 0$ for all $m \neq n$, only when the image pixel size Δx equals the Fourier pixel size $\Delta x_f = 1/(N_{data}\Delta k)$. This is because

$$\begin{aligned}E\{e(m\Delta x)e^*(n\Delta x)\} &= \Delta k^2 \sum_{p\in\mathcal{N}_{data}} \sum_{q\in\mathcal{N}_{data}} E\{\varepsilon(p\Delta k)\varepsilon^*(q\Delta k)\}e^{i2\pi(mp-nq)\Delta k\Delta x} \\ &= \Delta k^2 \sum_{p\in\mathcal{N}_{data}} E\{\varepsilon(p\Delta k)\varepsilon^*(p\Delta k)\}e^{i2\pi(m-n)p\Delta k\Delta x} \\ &= \Delta k^2\sigma^2_{data}\frac{\sin[\pi N_{data}(m-n)\Delta k\Delta x]}{\sin[\pi(m-n)\Delta k\Delta x)]}e^{-i\pi(m-n)\Delta k\Delta x}\end{aligned} \tag{21}$$

which is equal to zero for all $m \neq n$ only when $\Delta x = 1/(N_{data}\Delta k)$.

The average signal power of each pixel in the Fourier reconstruction is, in general, given by

$$\begin{aligned}S^2 &= \frac{1}{N_{imag}} \sum_{m\in\mathcal{N}_{imag}} \rho_{DFT}(m\Delta x)\rho^*_{DFT}(m\Delta x) \\ &= \frac{1}{N_{imag}}\Delta k^2 \sum_{m\in\mathcal{N}_{imag}} \sum_{n\in\mathcal{N}_{data}} \sum_{l\in\mathcal{N}_{data}} s(n\Delta k)s^*(l\Delta k)e^{i2\pi(n-l)m/N_{imag}} \\ &= \Delta k^2 \sum_{n\in\mathcal{N}_{data}} |s(n\Delta k)|^2.\end{aligned} \tag{22}$$

Hence, the average image SNR per pixel is

$$\text{SNR}|_{pixel} = \frac{S}{\sigma_{imag}} = \frac{\sqrt{\sum_{n\in\mathcal{N}_{data}} |s(n\Delta k)|^2}}{\sqrt{N_{data}}\sigma_{data}}. \tag{23}$$

The following conclusions can be drawn from the above discussions.

1. The average image SNR per pixel is independent of zero-padding. Therefore, reducing the image digital pixel size by zero-padding does not affect the image SNR. This is because zero-padding is no more than an interpolation method, albeit a higher-order one. The signal strength per pixel remains the same, which is the integration of the true image function over a Fourier pixel and the same is true for the noise. However, when zooming by a factor of M with zero padding, the image noise from pixel to pixel is not guaranteed to be uncorrelated. This means that one will not obtain a factor of $\sqrt{M}$ improvement in SNR by averaging M such pixels in this case.
2. Since $s(n\Delta k)$ decays as fast as $1/(n\Delta k)$, S is a very slowly growing function of N_{data}. Thus, if one attempts to improve image resolution by extending k-space coverage from $N_{data}\Delta k$ to $2N_{data}\Delta k$, the image SNR will suffer a loss by approximately a factor of $\sqrt{2}$ because of the increased noise included. This property is characteristic of the Fourier reconstruction method. With some modeling methods, to be discussed later, this weakness can be overcome.

To give an intuitive understanding of these important properties, let's consider a hypothetical example, shown in Fig. 3. Suppose that we have a 1-D object and four phase-encoding measurements. Direct application of the FFT to the four data points gives four pixels with SNR_0. If the data is zero-filled to eight points, the resultant eight pixels

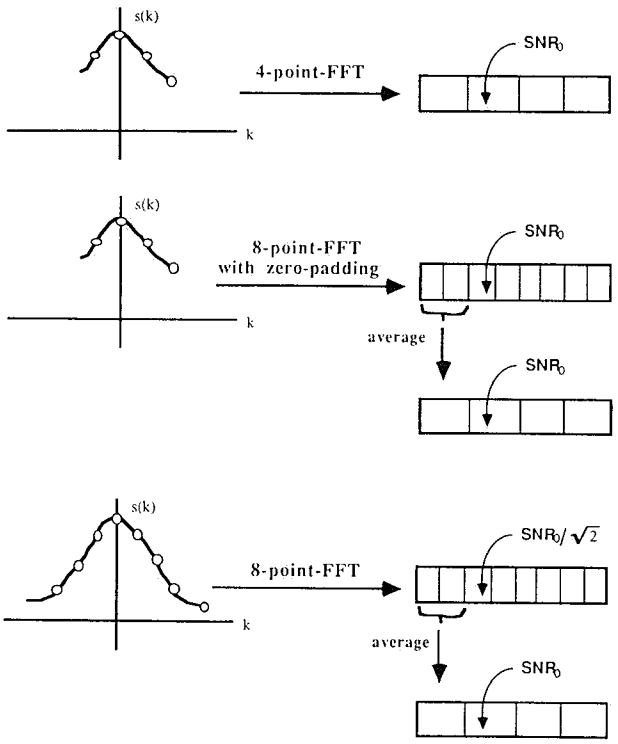

Figure 3. An illustration of the SNR efficiency for the Fourier method.

will have the same signal-to-noise ratio, SNR_0. Note that in both cases, the Fourier pixel is the same, and so is the signal strength per pixel. If we recollapse the eight pixels to four pixels, no SNR improvement will be gained, since the noise is correlated. On the other hand, if eight phase-encoded measurements were actually collected (for improved resolution), the new image signal-to-noise ratio, SNR_1, would be about $SNR_0/\sqrt{2}$ [actually, slightly higher because of the added signal energy from the new sample points; the exact value can be calculated from Eq. (23)]. Recollapsing the eight pixels to four will regain the original signal-to-noise ratio, but the imaging time is doubled; consequently, the SNR efficiency[5] is reduced by a factor of $\sqrt{2}$. Therefore, with the Fourier reconstruction methods, improving spatial resolution with extended k-space sampling will lead to an irreversible loss of SNR efficiency.[6] If a reconstruction method can give the same spatial resolution with a reduced set of phase-encoded measurements, it will improve the imaging SNR efficiency as well. This is, in fact, one of most important objectives to be achieved with modern constrained reconstruction methods.

Examples

To demonstrate the effects of data truncation (symmetric or asymmetric) and filtering, a set of reconstruction results is presented in Figs. 4–8. One can see from Fig. 4 that the

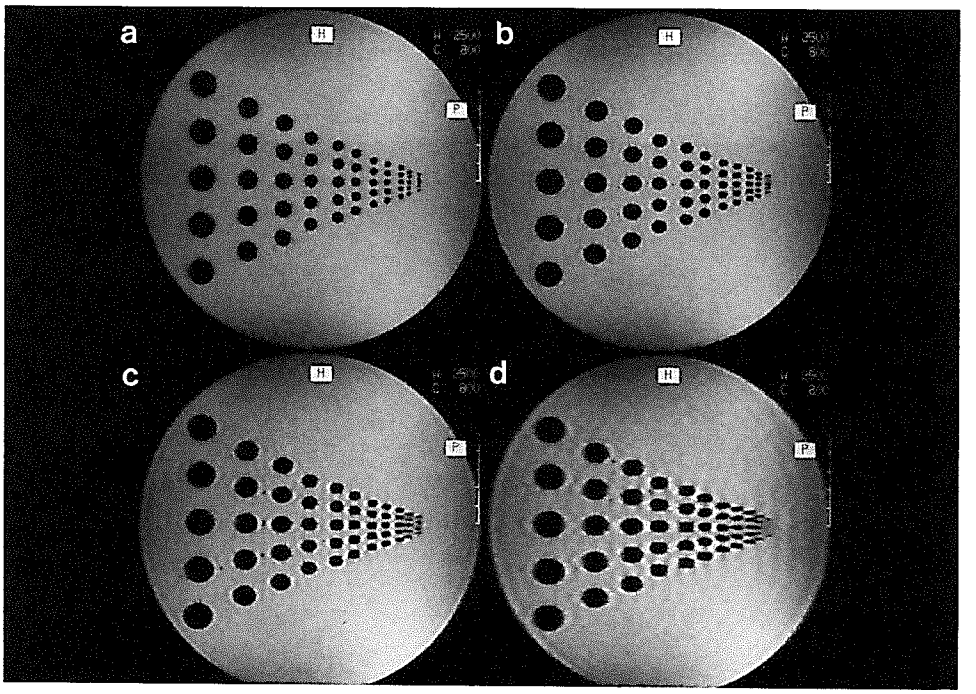

Figure 4. Fourier reconstructions of a phantom image. The vertical direction is reconstructed with 256 data points and the horizontal direction with (a) 256 data points; (b) 128 data points; (c) 96 data points; and (d) 64 data points. Note the increased Gibbs ringing and loss of resolution as the number of data points is reduced.

[5]SNR efficiency is defined as $SNR|_{pixel}/\sqrt{\text{total imaging time}}$.

[6]One is also referred to (62) for a similar argument.

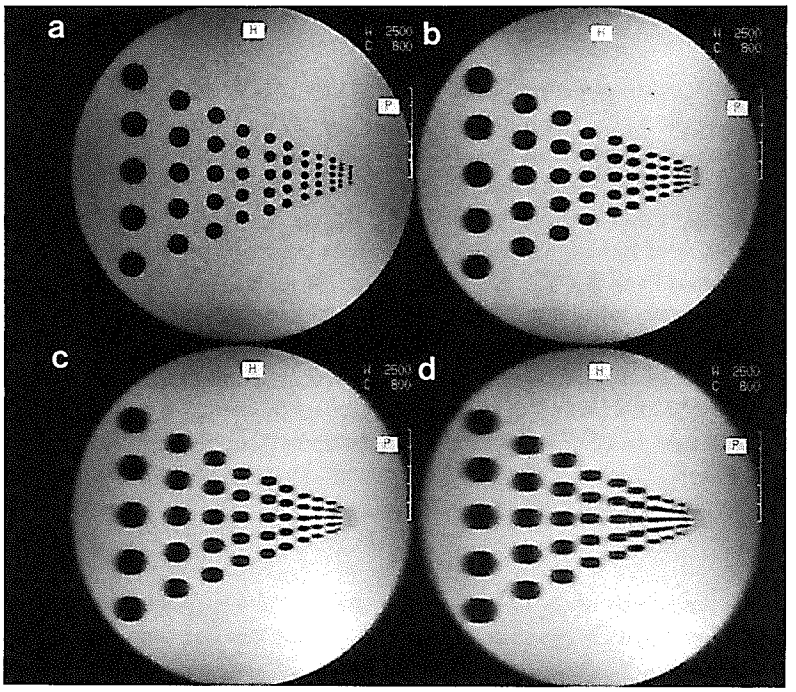

Figure 5. Hamming-filtered Fourier reconstructions corresponding to Fig. 4. Note that the Gibbs ringing in Fig. 4 is removed but with a further loss of image resolution.

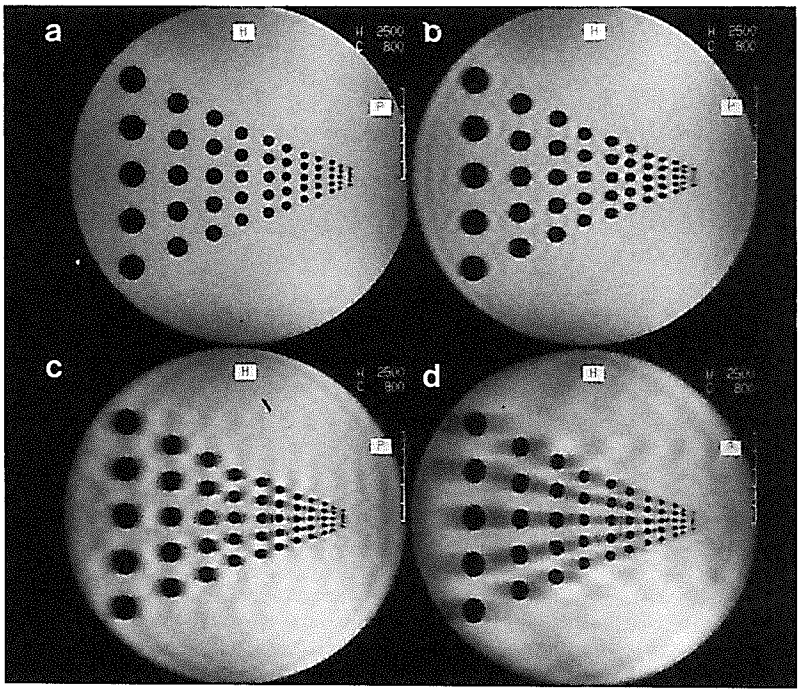

Figure 6. Fourier reconstruction from asymmetric data sets along the horizontal direction with different degrees of asymmetry: (b) 32 data points before the echo and 96 points after, or simply $32 \leqslant n < 96$; (c) $-16 \leqslant n < 112$; and (d) $-8 \leqslant n < 120$. Note that fixing the total number of data points, resolution improves as the data asymmetry increases, but the Gibbs ringing becomes more destructive.

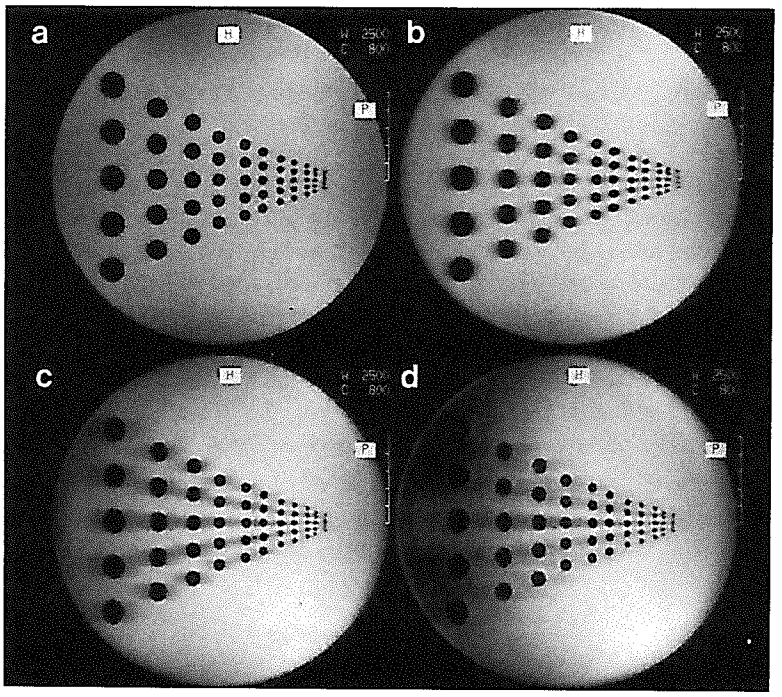

Figure 7. Hamming-filtered Fourier reconstructions corresponding to Fig. 6. The filter described in Eq. (11) is applied to the first n_0 data points before the echo. Note that the Gibbs ringing in Fig. 6 is effectively removed by this filter, but some "shadowing artifacts" and geometric distortions in the circular structures remain.

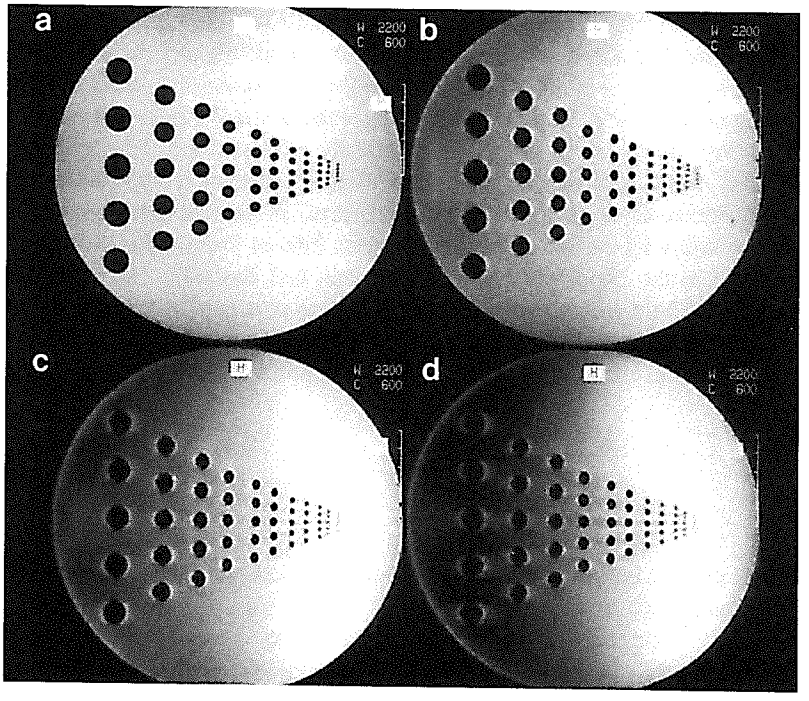

Figure 8. The same as Fig. 7, but the filter described in Eq. (12) is applied to the central $2n_0$ data points on both sides of the echo. Note the improved resolution and enhanced edge artifacts compared with the images in Fig. 7.

Gibbs ringing becomes progressively more serious as the number of data points is reduced from 256 points to 64 points in the horizontal direction. The loss of resolution is obvious in the regions of circular structures as fewer data points are used. The Hamming-filtered reconstructions in Fig. 5 remove the spurious ringing, but at the expense of already limited resolution. In the asymmetric sampling case, the Gibbs ringing comes dominantly from the truncation of one side of the data. Therefore, the ringing period in Fig. 6 is about $1/(2n_0\Delta k)$. When an asymmetric Hamming filter is applied to roll off the sharp transition in this side of the data, the Gibbs ringing is removed as shown in Fig. 7. Often the filter described in Eq. (12) is applied on both sides of the data for improved resolution, as proposed in (171). The disadvantages of this filter are increased edge artifacts (phase artifacts) and reduced SNR, as shown in Fig. 8. (Note, some of these artifacts are removed by the phase correction step in the Margosian algorithm, to be described in a later section).

PHASE-CONSTRAINED RECONSTRUCTION

Introduction

Asymmetric sampling in Fourier space arises in magnetic resonance imaging 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. The source of the data sampling asymmetry in gradient echo sequences can be due to gradient design, or it can occur in an expected symmetric echo when there is a field inhomogeneity present, which shifts the echo away from the expected echo time (84). It is also sometimes used in the phase-encoding direction when an asymmetric set of phase-encodings is applied to collect as many high frequency components as possible while reducing the data acquisition time. In the following discussion, we assume that $\hat{s}(k)$ is available at $k = n\Delta k$, for $n \in \mathcal{N}_{data} = \{-n_0, -n_0 + 1, \cdots, N - 1\}$, representing an asymmetric echo in either the phase-encoding or read-out direction. Usually, n_0 is much smaller than N, typically, $n_0 = 16$ or 32, but any values of n_0 can occur especially with gradient-echo imaging. Under this condition, direct application of the conventional Fourier reconstruction method would lead to serious asymmetric truncation artifacts, as demonstrated in Fig. 6. When the degree of sampling asymmetry is small, some filtered-Fourier method can be used (171), but always at the expense of image resolution and sharpness, as discussed previously. Arguably, the best possible reconstruction one can expect in this case is the one that would be obtained if the hypothetical symmetric data set were available. In other words, the goal of partial-Fourier reconstruction is to find an approximated image function $\hat{\rho}(x)$ such that

$$s(n\Delta k) = \sum_{m=-N}^{N-1} \hat{\rho}(m\Delta x)e^{-i\pi mn/N}, \qquad -n_0 \leq n < N. \tag{24}$$

When the true image function $\rho(x)$ is known to be a real function this problem is trivial, since the unknown data, $s(-n\Delta k)$ for $-N \leq n < -n_0$, can easily be generated from the corresponding known part based on the conjugate symmetric relationship for $s(n\Delta k)$, namely, $s(-n\Delta k) = s^*(n\Delta k)$. This technique known as the *conjugate symmetrization* method has been used by Feinberg et al. (64) with some success in spin-echo imaging

applications. The **cosine**-transform (31) can also be used to find $\hat{\rho}(m\Delta x)$ from $s(n\Delta x)$, for $n \geq 0$ only. However, $\rho(x)$ is in practice a complex function, since motion and magnetic field inhomogeneity often introduce a nonzero phase $\varphi(x)$ to $\rho(x)$. In this case, without additional constraints, $\hat{\rho}(m\Delta x)$ in Eq. (24) is underdetermined by the data, since there are $2N$ unknowns but only $(N + n_0)$ equations. In this chapter, we will review a specialized class of constrained methods developed for this problem. These methods basically follow a characteristic two-step procedure in which a phase image is first estimated and then it is used as a constraint to get the final magnitude reconstruction (hence, the name *phase-constrained reconstruction*). Other more general methods developed for finite-Fourier imaging discussed in later sections, can also be applied to this problem. Note that phase-constrained reconstruction can also be done using a brute-force matrix inversion method (79), which reformulates Eq. (24) as a matrix equation, incorporating the required phase constraint. Such an approach is usually not practically useful and will not be included for discussion here, since it is vulnerable to phase errors as well as computationally expensive, although regularization schemes can be applied to stabilize the solution (26,73). The following discussion will be focused on how to perform the phase-constrained reconstruction without using the matrix approach.

Margosian Method

Theory. The Margosian method represents one of the earlier but rather successful methods for partial-Fourier reconstruction in spin-echo imaging applications (156). This method has also been recently discussed by Macovski (173) based on a homodyne detection theory. The central idea of the Margosian method lies in the observation that the Fourier transform reconstruction from a half echo, $\rho_{\text{half}}(x)$, is related to the true image function $\rho(x)$ by the following convolution (in the continuous case):

$$\rho_{\text{half}}(x) = \rho(x) \bigstar 0.5\left[\delta(x) + i\frac{1}{\pi x}\right] = 0.5|\rho(x)|e^{i\varphi(x)} + i0.5[|\rho(x)|e^{i\varphi(x)}] \bigstar \frac{1}{\pi x}. \quad (25)$$

This is derived by considering $s(k)$ being filtered by the Heaviside window function

$$w(k) = \begin{cases} 0 & \text{for } k < 0 \\ 1 & \text{for } k \geq 0. \end{cases} \quad (26)$$

Phase-correcting $\rho_{\text{half}}(x)$ with an estimated phase $\hat{\varphi}(x)$ gives

$$\rho_{\text{half}}(x)e^{-i\hat{\varphi}(x)} = 0.5|\rho(x)|e^{i[\varphi(x)-\hat{\varphi}(x)]} + i0.5\left\{[|\rho(x)|e^{i\varphi(x)}] \bigstar \frac{1}{\pi x}\right\} e^{-i\hat{\varphi}(x)}. \quad (27)$$

When $\hat{\varphi}(x)$ is close to $\varphi(x)$, the first term gives the desired magnitude reconstruction.

To find the phase constraints $\hat{\varphi}(x)$, Margosian, Schmitt, and Purdy (156) proposed to apply the Fourier transform to the central, low-frequency data $[s(n\Delta k), -n_0 \leq n \leq n_0 - 1]$, assuming that only low-frequency phase variations exist in the image function. This reconstructs a $2n_0$-point resolution phase image as

$$\varphi_{\mathrm{DFT}}(x) = \arg\left\{\sum_{n=-n_0}^{n_0-1} s(n\Delta k)e^{i2\pi n\Delta k x}\right\}. \tag{28}$$

Note that, if all the available data points were used for the phase calculation, $\varphi_{\mathrm{DFT}}(x)$ in Eq. (28) would be of $(N + n_0)-$point resolution. But these phase constraints are not useful for use in Eq. (27), since $\varphi_{\mathrm{DFT}}(x) = \arg\{\rho_{\mathrm{half}}(x)\}$ and the Margosian reconstruction will become identical to the truncated Fourier reconstruction from a half echo with its associated artifacts.

Algorithm. In summary, the Margosian method for image reconstruction from partial-Fourier data consists of the following steps.

1. *Phase estimation:* The central low frequency data, $s(n\Delta k)$ for $-n_0 \leq n < n_0$, are zero-filled to $\pm N$ and Hamming-filtered to roll off the sharp transitions between the measured data and the padded zeros to reduce Gibbs' ringing. The FFT is then applied and $\varphi_{\mathrm{DFT}}(x)$ is obtained for $x = n/(2N\Delta k)$, $-N \leq n < N$ to provide $2N$ phase constraints for pixel-by-pixel phase correction.
2. *Filtered Fourier reconstruction:* Instead of obtaining the Fourier reconstruction from the half echos, $\rho_{\mathrm{half}}(x)$, as described in the previous *theory* subsection, the first n_0 data points before the echo are also used in practice to avoid the discontinuity at the data origin. This asymmetric data set is zero-filled to $-N$ and the central $2n_0$ data points are filtered with an asymmetric Hamming filter $[w(n) = 0.5 + 0.5 \cos [\pi(n - n_0)/(2n_0)]; -n_0 \leq n < n_0]$ to roll off the sharp transition between the zeros and the measured data. The FFT is then applied to this zero-padded, filtered data set to get an approximated half-echo reconstruction, $\hat{\rho}_{\mathrm{half}}(x)$.
3. *Phase correction:* $\hat{\rho}_{\mathrm{half}}(x)$ is phase-corrected pixel by pixel using the phase obtained from Step 1, and the real part is then extracted from it as the final reconstruction.

Characteristics. The properties of the Margosian method are summarized as follows.

- It is simple and can handle small phase variations that occur in spin-echo imaging applications.
- It is, however, very sensitive to phase estimation errors. Notably, the reconstruction is proportional to[7]

$$\cos[\varphi(x) - \varphi_{\mathrm{DFT}}(x)]|\rho(x)| \tag{29}$$

and the **cos**-modulation can be significant in regions where large phase errors exist. Additionally, when large image phase variations exist, such as in the case of gradient-echo imaging, this method cannot guarantee the right reconstruction even if the exact phase constraint is available. This is because the real part of the second term in Eq. (27), which equals

$$\left\{[|\rho(x)| \cos \varphi(x)] \bigstar \frac{1}{2\pi x}\right\} \sin \varphi_{\mathrm{DFT}}(x) - \left\{[|\rho(x)| \sin \varphi(x)] \bigstar \frac{1}{2\pi x}\right\} \cos \varphi_{\mathrm{DFT}}(x), \tag{30}$$

[7]See the first term of Eq. (27).

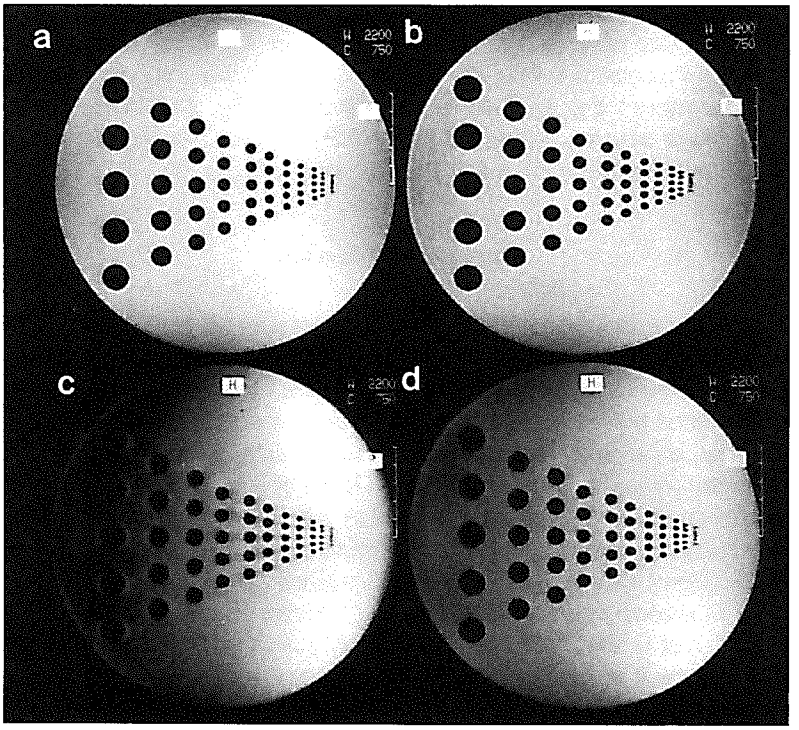

Figure 9. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a) the Fourier method for full symmetric echos of 256 data points, (b) the zero-padded Fourier method from truncated symmetric echos of 128 data points; and the Margosian method from 128 asymmetric data points with different levels of sampling asymmetry: (c) $-16 \leqslant n < 112$; and (d) $-32 \leqslant n < 96$. Note the improved resolution of the Margosian reconstructions over the Fourier reconstruction from 128 symmetric data points but also the artifacts compared with the full-echo reconstruction.

is significant when $\varphi(x)$ as well as $\varphi_{\mathrm{DFT}}(x)$ is not small. This term can contribute serious artifacts in the form of geometric distortions to the final image because of the convolution with the filter function $1/(2\pi x)$ and the **sin-** and **cos**-modulations.

Figure 9 shows a set of reconstructions from a phantom data set acquired using a FISP sequence. The imaging parameters were $T_R = 45ms$ and $T_E = 12ms$ and the matrix size of the raw data was 256×256. Reconstruction along the vertical (phase-encoding) direction was first performed using the standard Fourier method on the full echos, but the horizontal direction was reconstructed with different levels of truncation. For a comparison, the reconstructions obtained using the conventional Fourier method from symmetric echos of 256 and 128 data points, respectively, are presented in Figs. 9a–b, and the Margosian reconstructions from 128 asymmetric data points are presented in Figs. 9c–d. Different levels of sampling asymmetry were used in the Margosian reconstruction to illustrate its sensitivity to the quality of the phase constraints. Also, the number of data points on the right hand side of the echo was chosen so that the matrix size of the truncated

data sets was a constant (256 × 128).[8] The image in Fig. 9c was reconstructed using a 32 point resolution phase estimate ($n_0 = 16$ points before the echo). Note the improved resolution over the reconstruction in Fig. 9b but also "shadowing artifact" because of the phase variations across the image (see the phase image in Fig. A1). When the number of points for the phase estimate is doubled ($n_0 = 32$), Fig. 9d, this artifact is reduced but remains significant because the term in Eq. (30) is non-negligible.

Cuppen-POCS Method

Theory. To overcome the problems with the Margosian method, Cuppen and Van Est (48) proposed an iterative extrapolation approach to generate a full symmetric data set for use in the normal Fourier reconstruction method. Unlike the earlier conjugate symmetrization method, this method utilizes a pre-estimated phase to constrain the symmetrization process. In this way, if the exact phase constraints are available, one can potentially recover the full symmetric data set. On the other hand, if the phase constraints are not accurate, exaggeration of the phase errors can be avoided by terminating the iterative process early. The phase constraints commonly used are obtained from the Fourier transform of the central low frequency data as in the Margosian method. The phase-constrained symmetrization step is accomplished originally by an iterative method that forces the phase constraints such that in the l^{th} iteration, $\hat{\rho}_l(x) = \hat{\rho}^*_{l-1}(x)e^{i2\hat{\varphi}(x)}$; the image thus obtained is then transformed back to create a pseudo data set $\bar{s}_l(k)$ to provide data points for the unsampled data region. This procedure, however, may not converge to a symmetrized data set with the required phase since $\hat{\rho}_l(x)$ is not constrained to have the required phase $\hat{\varphi}(x)$. To overcome this problem, Liang (134) suggested the use of a more general mathematical framework called projection onto convex sets (POCS) (207,208,226,254,255). This method was further studied by Mitchell (168), Lindskog (147), and Haacke (85,86).

In this formalism, one first conceptually constructs the following two convex sets:

$$\Omega_1 = \{\hat{\rho}(x) | \arg\{\hat{\rho}(x)\} = \hat{\varphi}(m\Delta x) \text{ for } m \in \mathcal{N}_{imag}\},$$
$$\Omega_2 = \{\hat{\rho}(x) | \mathcal{F}\{\hat{\rho}(x)\} = s(n\Delta k) \text{ for } n \in \mathcal{N}_{data}\}.$$

Namely, Ω_1 contains all the images satisfying the predetermined phase constraints and Ω_2 contains all the images consistent with the measured data. The phase-constrained problem is then to find an image function $\hat{\rho}(x)$ in the intersection of Ω_1 and Ω_2, i.e., $\hat{\rho}(x) \in \Omega = \Omega_1 \cap \Omega_2$. When Ω is not empty (i.e., the phase constraints are not poorly chosen), such an image function exists and can be found using the POCS theory by alternating projections of an initial estimate onto these two sets.

The following mappings, $\wp_1$ and $\wp_2$, project any image function $\hat{\rho}(x)$ onto Ω_1 and Ω_2, respectively,

$$\wp_1\{\hat{\rho}(x)\} = |\hat{\rho}(x)|e^{i\hat{\varphi}(x)};$$
$$\wp_2\{\hat{\rho}(x)\} = \mathcal{F}^{-1}\mathcal{R}\mathcal{F}\{\hat{\rho}(x)\},$$

[8] This was to keep the total imaging time constant if partial-Fourier imaging is used in the phase-encoding direction. In many practical applications in the readout direction, one can always collect a complete half echo, i.e., $-n_0 \leq n < N$, instead of $-n_0 \leq n < N - n_0$, without increasing the imaging time, but the conclusions here about the behavior of the Margosian method hold for both situations.

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the FFT operator and its inversion, and $\mathcal{R}$ is the "data replacement" operator defined as

$$\mathcal{R}\{\tilde{s}(n\Delta k)\} = \begin{cases} s(n\Delta k) & -n_0 \leq n \leq N-1 \\ \tilde{s}(n\Delta k) & \text{otherwise.} \end{cases}$$

The desired reconstruction, $\hat{\rho}(x)$ can then be obtained from the following POCS iteration

$$\hat{\rho}_{l+1}(x) = \wp_1\wp_2\{\hat{\rho}_l(x)\}, \tag{31}$$

where the initial condition $\hat{\rho}_0(x)$ can be chosen to be the zero-filled Fourier reconstruction.

Algorithm. The Cuppen-POCS method for image reconstruction from partial-Fourier data consists of the following steps.

1. As in the Margosian method, the phase constraints, $\varphi_{\text{DFT}}(x)$, are obtained from the central low frequency data [$s(n\Delta k)$ for $-n_0 \leq n < n_0$] by a zero-filled, filtered Fourier transformation.
2. Zero-fill the asymmetric data to $-N$ and apply the inverse FFT to get a first approximation, $\hat{\rho}_0(x)$, of the required reconstruction $\hat{\rho}(x)$.
3. In the lth iteration, force the previously estimated image, $\hat{\rho}_{l-1}(x)$, to the required phase $\varphi_{\text{DFT}}(x)$ using

$$\hat{\rho}_l(x) = |\hat{\rho}_{l-1}(x)|e^{i\varphi_{\text{DFT}}(x)}. \tag{32}$$

4. Apply the FFT to $\hat{\rho}(x)$ to get a pseudo data set $\tilde{s}_l(n\Delta k)$ and then fill the unsampled data region ($-N \leq n < n_0$) of the measured data $s(n\Delta k)$ with $\tilde{s}_l(n\Delta k)$. In the final iteration, $\tilde{s}_l(n\Delta k)$ and $s(n\Delta k)$ often need to be properly merged at their boundary ($n = -n_0$) to remove a possible discontinuity between the original and extrapolated data and thus to avoid Gibbs ringing. The merging is often done using a linear weighting scheme, such as

$$s_{\text{merge}}[(-n_0 + m)\Delta k] = \frac{M-m}{M}\tilde{s}_l[(-n_0 + m)\Delta k] + \frac{m}{M}s[(-n_0 + m)\Delta k], \tag{33}$$

   where $0 \leq m < M$ with M being the number of data points to be merged.
5. Apply the inverse FFT to the extrapolated (symmetrized) data set to get a new reconstruction and go back to step 3 until the maximum number of iterations is reached.

Characteristics. The properties of the Cuppen-POCS method are summarized as follows.

- It is flexible and can avoid exaggeration of phase errors by early termination of the iteration.
- Its convergence behavior is well understood. The composite operator $\wp = \wp_1\wp_2$ is non-expansive in general, i.e.,

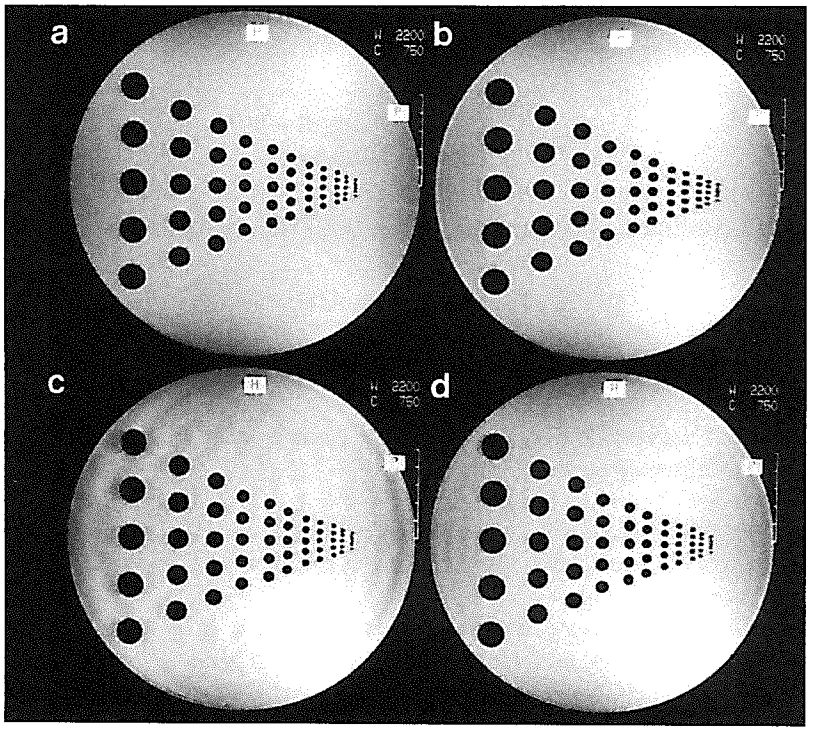

Figure 10. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a) the Fourier method from full symmetric echos of 256 data points; (b) the zero-padded Fourier method from truncated symmetric echos of 128 data points; and the Cuppen-POCS method with 128 asymmetric data points with different levels of sampling asymmetry: (c) $-16 \leq n < 112$; and (d) $-32 \leq n < 96$. Note the ringing artifacts in the POCS reconstruction (c) when the phase estimate is poor.

$$\|\wp\{\rho_1(m\Delta x)\} - \wp\{\rho_2(m\Delta x)\}\| \leq \eta\|\rho_1(m\Delta x) - \rho_2(m\Delta x)\| \tag{34}$$

where $0 \leq \eta \leq 1$. But it behaves "almost" like a contractive operator[9], since for practical images the chance that $\eta = 1$ is very small. Therefore, in practice, the POCS iteration can converge rapidly (exponentially), since

$$\|\hat{\rho}_l(m\Delta x) - \hat{\rho}(m\Delta x)\| < \frac{\eta^l}{1-\eta}\|\hat{\rho}_l(m\Delta x) - \hat{\rho}_0(m\Delta x)\|, \tag{35}$$

where $\hat{\rho}_0(m\Delta x)$ is the initial condition and $\hat{\rho}(m\Delta x)$ is the required solution. Usually, 2 to 4 iterations are sufficient.

- When the phase estimate is poor, however, the Cuppen-POCS reconstruction may suffer from phase errors in the form of signal cancellation and spurious ringing.

[9]See the Appendix for definition. For detailed discussion of contractive mapping and its applications, see (114,133,203,210,233,234,235).

Artifacts in the regions where high spatial frequency phase behavior exists can sometimes be reduced by applying the phase constraints only outside these regions (86).

Figure 10 shows a set of reconstructions from the same data set used in Fig. 9 but using the Cuppen-POCS algorithm. Organized as in Fig. 9, the images in Figs. 10c–d were reconstructed using 32 ($n_0 = 16$) and 64 ($n_0 = 32$) point resolution phase estimates, respectively. As can be seen, when the phase estimate is poor, the Cuppen-POCS reconstruction, Fig. 10c, shows some noticeable ringing artifacts. These artifacts come from the inconsistency between the extrapolated data and the measured data and display the same periodicity as in the zero-padded Fourier reconstruction in Fig. 6. But unlike that in the Margosian reconstructions, this artifact is significantly reduced in the Cuppen-POCS reconstruction when 64 data points ($n_0 = 32$) are used for phase estimation, as shown in Fig. 10d. This is an apparent advantage of the Cuppen-POCS method over the Margosian method, which is only useful when $\varphi(x)$ as well as $|\varphi(x) - \hat{\varphi}(x)|$ is small, as discussed in the previous section. In gradient-echo imaging experiments, $\varphi(x)$ is often non-negligible, but the Cuppen-POCS method will always yield a good reconstruction as long as the phase estimation is reasonable, i.e., $|\varphi(x) - \hat{\varphi}(x)|$ is small.

Summary

Magnetic resonance imaging methods usually lead to the collection of complex Fourier transform data. If collected as an echo, the data set contains both positive and negative spatial frequency components which, for a real image function $\rho(x)$, obey a simple Hermitian symmetry relationship. This symmetry allows one half the data to be collected to reconstruct the same image with only $\sqrt{2}$ loss in SNR. The advantage of acquiring the data in this fashion is either a reduction in imaging time when applied in the usual phase-encoding direction or a reduction in T_E when applied in the read-out direction. Unfortunately, the realness of $\rho(x)$ is destroyed by certain features of the data acquisition sequences that allow local phase errors to persist; hence, the Hermitian symmetry is violated. The resulting artifacts in the image when these phase errors are not accounted for are usually unacceptable. This is especially true for gradient-echo images where field inhomogeneities play a key role in determining local phase errors. Similar phase errors can also occur in both gradient-echo and spin-echo images in regions of moving spins when velocity is uncompensated.

All of the existing partial-Fourier imaging methods take n_0 "extra" points in the low spatial frequency domain and usually the central $2n_0$ points can provide a reasonable phase estimate[10]. However, if an echo is unexpectedly shifted outside the sampling window, none of the methods described in this section will work. In practice, one needs to choose n_0 carefully to make sure that the echo is captured, which requires that the phase per pixel be less than $(n_0/N)\pi$. For a given field inhomogeneity $\Delta B(x)$, the maximum phase error is known to be $\gamma \Delta B(x) \Delta x T_E$ and n_0 can be calculated accordingly for a given situation. Once n_0 is known, the loss of image SNR can be calculated as compared to the ideal hypothetical symmetrical sampling case. Usually, a loss of $\sqrt{2N/(N + n_0)}$ is expected, but this loss is usually well compensated for by an improved contrast-to-noise

[10]Some methods discussed in later sections use no phase constraints.

ratio in many applications, such as angiographic imaging. A summary of the salient features of the discussed methods for asymmetric Fourier reconstruction is given as follows.

1. **Conventional FFT or filtered FFT method**
   - *Advantages:*
     — very efficient.
   - *Disadvantages:*
     — image blurring and loss of resolution.
2. **Margosian method**
   - *Advantages:*
     — improved resolution;
     — efficient;
     — useful when small phase variations are present (e.g., spin-echo imaging).
   - *Disadvantages:*
     — cannot handle large phase variations (leading to image distortions);
     — SNR loss beyond the usual factor of $\sqrt{2}$.
3. **Cuppen-POCS method**
   - *Advantages:*
     — improved resolution;
     — fairly efficient;
     — useful for both spin-echo and gradient-echo imaging applications.
   - *Disadvantages:*
     — cannot handle high degrees of data asymmetry;
     — introduces phase errors in the form of spurious ringing and signal cancellations;
     — SNR loss by a factor of $\sqrt{2}$.

NONPARAMETRIC CONSTRAINED METHODS

Introduction

We now return to discussing finite-Fourier imaging in general. Recall from the first section that in finite-Fourier imaging, we collect a finite set of data points, $s(n\Delta k)$ for $n \in \mathcal{N}_{data}$, with either symmetric or asymmetric sampling in k-space. The goal of image reconstruction is to find an image function $\hat{\rho}(x)$ as close to the true image function $\rho(x)$ as possible. In this section, we will review the class of nonparametric methods proposed for this problem. Common to all these methods is that the constraints used permit the use of the conventional Fourier series representation for the image function, as opposed to the parametric methods, to be discussed in the next section, which represent the image function in terms of a parametric model. These constraints bring in new information, in one way or another, by recovering some of the unavailable high frequency components, and the conventional Fourier reconstruction technique is then used to generate the final image, as shown below in Fig. 11. How accurately the missing high frequency data are recovered depends on the constraints themselves, and the accuracy of the extrapolated data determines the quality of the final reconstruction. In this section, we will review extrapolation methods using constraints such as *finite spatial support, maximum entropy, linear predictability,* and *image smoothness.*

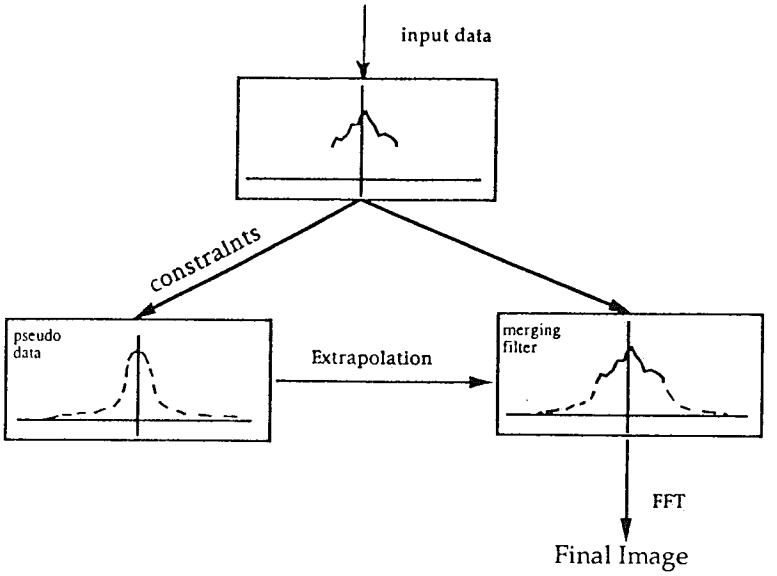

Figure 11. Block diagram of nonparametric constrained reconstruction methods.

Support-Limited Extrapolation

Theory. In any medical imaging problem, the image function $\rho(x)$ is always support-limited (SL) (see first section for definition). It is well known that the Fourier transform $s(k)$ of such a function is analytic over the entire real axis and can be recovered, in principle, from any finite segment of it. Based on this fact, several celebrated continuous extrapolation algorithms have been developed, such as the methods using the prolate spheroidal wave function decomposition theory by Slepian et al. (211,212) and the iterative support-limited extrapolation method independently developed by Gerchberg (67) and Papoulis (175). These method have enjoyed great success in restoring optical images from diffraction-limited systems. For the problem we are considering here, however, there is only a finite set of sampled data points available. In this discrete case, the analytic property of $s(k)$ becomes meaningless, but the basic idea of making use of the support-bound information is still appropriate and has been thoroughly investigated by several research groups; Refer to (40,59,103,193,195–200,203) for detailed discussions.

Assuming that $\rho(x)$ vanishes outside of $\mathcal{B} = (-L_x/2, L_x/2)$, we have

$$\rho(x) = w(x) \sum_{m=-\infty}^{\infty} a(m\Delta k) e^{i2\pi m\Delta kx} \tag{36}$$

where $w(x)$ is a rectangular window function defined as

$$w(x) = \begin{cases} 1, & |x| \leq L_x/2; \\ 0, & \text{otherwise} \end{cases} \tag{37}$$

and $a(m\Delta k)$ is an appropriate set of coefficients that are related to the measured data $s(k)$ by

$$s(k) = \sum_{m=-\infty}^{\infty} a(m\Delta k) \frac{\sin[\pi L_x(k - m\Delta k)]}{\pi(k - m\Delta k)} . \tag{38}$$

This is the well-known *Shannon interpolation formula* (176). Since $s(k)$ is only known for $k = n\Delta k$, $n \in \mathcal{N}_{data}$, for a finite $\mathcal{N}_{data}$, the nonuniqueness of the solution for the $a(m\Delta k)$ is evident. Therefore, unlike the continuous extrapolation problem, discrete extrapolation is inherently underdetermined. If the minimum-norm solution is sought, namely, $\Sigma_m |a(m\Delta k)|^2$ is minimized, one will find that $a(m\Delta k)$ satisfies the following conditions

$$\begin{cases} a(m\Delta k) = 0, & m \notin \mathcal{N}_{data}; \\ s(n\Delta k) = \displaystyle\sum_{m \in \mathcal{N}_{data}} a(m\Delta k) \frac{\sin[\pi L_x(m - n)\Delta k]}{\pi(m - n)\Delta k} , & n \in \mathcal{N}_{data}. \end{cases} \tag{39}$$

Therefore, Eqs. (36) and (39) provide the governing equations for the SL-extrapolation reconstruction problem.

Algorithm. In practice, SL-extrapolation reconstruction can be obtained iteratively or non-iteratively using the following algorithms.

Non-iterative algorithm — In matrix form, Eq. (39) can be put as

$$\boldsymbol{E}\boldsymbol{a} = \boldsymbol{s} \tag{40}$$

where $\boldsymbol{E}$ is an $N_{data} \times N_{data}$ real matrix; $\boldsymbol{a}$ and $\boldsymbol{s}$ are N_{data}-element complex vectors with

$$\begin{aligned} a_m &= a(m\Delta k), \qquad m \in \mathcal{N}_{data} \\ s_m &= s(m\Delta k), \qquad m \in \mathcal{N}_{data} \\ E_{mn} &= \frac{\sin[\pi L_x(m - n)\Delta k]}{\pi(m - n)\Delta k} , \qquad m, n = 1, 2, \cdots, N_{data}. \end{aligned}$$

It has been proven (103) that the extrapolation matrix $\boldsymbol{E}$ has N_{data} positive eigenvalues such that $1 > \lambda_1 > \lambda_2 > \cdots > \lambda_{N_{data}} > 0$. The first $N_{data}L_x\Delta k$ eigenvalues are very close to 1, and the others gradually decay to zero. Since $\boldsymbol{E}$ is positive-definite, $\boldsymbol{a}$ is uniquely determined from Eq. (40). The major problem with this non-iterative solution lies in its sensitivity to noise, particularly to broadband noise because the extrapolation matrix $\boldsymbol{E}$ is usually ill-conditioned. Paradoxically, the smaller the sampling rate Δk with respect to the Nyquist rate $1/L_x$, the more information the support-bound contains, but the more ill-conditioned the extrapolation matrix $\boldsymbol{E}$ will be. The usual way to cope with this ill-conditioned problem is to regularize Eq. (40) by perturbing $\boldsymbol{E}$ with a diagonal matrix $\mu\boldsymbol{I}$. In this case, Eq. (40) becomes

$$(\boldsymbol{E} + \mu \boldsymbol{I})\boldsymbol{a} = \boldsymbol{s}. \tag{41}$$

Other advanced regularization methods are available to replace the diagonal matrix by a more complex regularization matrix incorporating new constraints (1,25,26,59,201). In summary, the regularized minimum-norm solution for the discrete extrapolation problem can be obtained in the following two steps:

1. Solve Eq. (40) or Eq. (41) for $\boldsymbol{a}$; and
2. Substitute $\boldsymbol{a}$ into Eq. (36) to get $\rho(x)$ or into Eq. (38) to get the extrapolated signal $s(k)$, followed by the conventional Fourier reconstruction.

Iterative algorithm — The popular iterative algorithm for SL extrapolation developed independently by Papoulis and Gerchberg is summarized as follows. One is referred to (176) for more details.

1. Zero-fill the available data to $\pm N_{imag}$ points and apply the inverse FFT to get a first approximation, $\hat{\rho}_0(x)$, of the required reconstruction $\hat{\rho}(x)$.
2. In the lth iteration, force the previously estimated image, $\hat{\rho}_{l-1}(x)$, to have the required spatial support bound by multiplying it with the window function defined in Eq. (37)

$$\tilde{\rho}_l(x) = \hat{\rho}_{l-1}(x)w(x). \tag{42}$$

3. Apply the FFT to $\tilde{\rho}_l(x)$ to get a new pseudo data set $\tilde{s}_l(n\Delta k)$ and then fill the unsampled data region of the measured data $s(n\Delta k)$ with $\tilde{s}_l(n\Delta k)$.
4. Fourier transform the extrapolated data set to get a new reconstruction, $\hat{\rho}_l(x)$ and go back to step 2 if the maximum number of iterations is not reached; otherwise, accept it as the final reconstruction. To avoid noise amplification, this iterative process is often terminated after a few iterations, depending on the data signal-to-noise ratio.

Characteristics. The quality of a reconstruction based on support-limited extrapolation relies on the amount of information contained in the superimposed spatial support bound. If the data happen to be over-sampled (i.e., the field of view is larger than the support-bound), this constraint can help to recover some of the unmeasured high-frequency data. Another advantage of the method is its computational efficiency. The iterative solution can be obtained with a few iterations (mainly Fourier transformations), and the non-iterative solution can be obtained efficiently using the Levinson algorithm (122).

One drawback of this method is that it is sensitive to noise and requires relatively high data signal-to-noise to assure a reasonable gain in resolution. More importantly, if the data is sampled at the Nyquist rate, i.e., $\Delta k = 1/L_x$ as is often the case in MRI, the support-bound constraint becomes trivial ($\boldsymbol{E} = L_x\boldsymbol{I}$ and $\boldsymbol{a} = \Delta k\boldsymbol{s}$) and no data extrapolation will be gained. Therefore, the SL extrapolation method is not generally useful for overcoming the truncation artifacts, but has found applications in other MRI problems (102,250,251). We include its discussion here mainly because of its importance in the field of constrained reconstruction.

Maximum Entropy Methods

The maximum entropy principle has its origin in information theory largely due to Shannon. Its applications to data processing and image reconstruction have been advo-

cated by many noted scientists. Although it has been claimed that maximum entropy methods (MEM) have some potential for removal of random noise and systematic artifacts, such as the Gibbs ringing, their usefulness for MR applications, if any, has not been adequately demonstrated or justified. Considering that there exists a vast body of literature discussing these methods (46,47,57,61,63,74,75,107,108,111,112,116,120,146, 160,169,177,186,240,259), we will give only a brief description of the principle and its properties with respect to MRI applications in this section.

Entropy measures. The core of MEM is the definition of an entropy measure, which plays the role of guiding the selection of a reconstruction from the infinitely many compatible with the experimental data. Since the classical least squares principle is insensitive to spurious ringings[11] and gives the Fourier series as the optimal reconstruction, it is hoped that the maximum entropy principle would overcome this problem by selecting an artifact-free reconstruction. There are several forms of entropy measures currently in use. Among them, the following definitions are the best known.

Shannon entropy — For a positive image with pixel values $\rho(l\Delta x)$, for $l \in \mathcal{N}_{imag}$, the Shannon entropy is defined as

$$H_S = -\sum_{l \in \mathcal{N}_{imag}} \rho(l\Delta x) \log \rho(l\Delta x). \tag{43}$$

For complex image functions, the following modified Shannon entropy measures are often used:

$$H_S = -\sum_{l \in \mathcal{N}_{imag}} [\rho^R(l\Delta x) \log \rho^R(l\Delta x) + \rho^I(l\Delta x) \log \rho^I(l\Delta x)], \tag{44}$$

where the real part $\rho^R(l\Delta x)$ and the imaginary part $\rho^I(l\Delta x)$ of the image function $\rho(l\Delta x)$ can be considered as two positive and independent quantities, as in microwave tomography (15), or

$$H_S = -\sum_{l \in \mathcal{N}_{imag}} |\rho(l\Delta x)| \log |\rho(l\Delta x)| \tag{45}$$

when one is only interested in the modulus of $\rho(l\Delta x)$, as is sometimes the case in MRI.

When an *a priori* guess (estimate) $\rho^0(l\Delta x)$ of $\rho(l\Delta x)$ is available, this constraint can be incorporated into the entropy expression as

$$H_S = -\sum_{l \in \mathcal{N}_{imag}} \rho(l\Delta x) \log \frac{\rho(l\Delta x)}{\rho^0(l\Delta x)}. \tag{46}$$

[11]This is evident by the fact that the LS error from addition of a high-frequency component to the image can be made arbitrarily small by increasing the frequency of oscillations (Riemann-Lebesque lemma).

This measure is often called the cross-entropy (209).

Burg entropy — Another popular entropy definition due to Burg (39) expresses the entropy as

$$H_B = \int_{\mathcal{B}} \log \rho(x) dx.$$

This entropy measure has found wide application in high resolution spectral estimation from short data records (116,186,240).

Maximum entropy reconstruction. Maximum entropy reconstruction is obtained by maximizing H_S or H_B under the data consistency constraint. In maximizing the Shannon entropy H_S, the data consistency constraint is expressed in terms of the following so-called chi-squared statistical test as

$$\chi^2 = \sum_{n \in \mathcal{N}_{data}} \left| s(n\Delta k) - \Delta x \sum_{l \in \mathcal{N}_{imag}} \rho(l\Delta x) e^{-i2\pi l \Delta x n \Delta k} \right|^2 \Big/ \sigma^2_{data}, \tag{47}$$

which measures the misfit of the reconstruction $\rho(l\Delta x)$ to the known data $s(n\Delta k)$ and σ_{data} is the standard deviation of the noise in the data domain. This constrained problem is often transformed to an unconstrained problem through the introduction of a Lagrange multiplier μ to form the following objective function

$$Q = H_S - \mu\chi^2. \tag{48}$$

The solution we seek corresponds to a critical point of Q, with the value of μ chosen so that $\chi^2 = N_{data}$. There is no known formal solution to this problem in general and iterative optimization algorithms are often employed to find an asymptotic solution. But, due to the highly nonlinear nature of the problem, the computational burden can sometimes be heavy. Efforts have been made to develop new algorithms for improved efficiency (227,260).

Maximization of Burg's entropy H_B under the data constraints that

$$s(n\Delta k) = \int_{\mathcal{B}} \rho(x) e^{-i2\pi n \Delta k x} dx, \qquad n \in \mathcal{N}_{data} \tag{49}$$

is, surprisingly, simpler. It leads to a closed-form solution for $\rho(x)$ as

$$\rho(x) = 1 \Big/ \left| h_0 + \sum_{m=1}^{N_{data}/2-1} h_m e^{i2\pi m \Delta k x} \right|^2, \tag{50}$$

where the coefficients are determined by the Toeplitz system

$$\sum_{m=0}^{N_{data}/2-1} s[(n-m)\Delta k] h_m = \begin{cases} \sigma, & n = 0; \\ 0, & n = 1, \ldots, N_{data}/2 - 1, \end{cases} \tag{51}$$

with σ being the noise standard deviation.

Characteristics. The Burg maximum entropy reconstruction is easy to calculate and has proved very useful for reconstructing a δ-function like image function with superresolution. But it is known that H_B is not an appropriate "configurational" entropy measure for general MR anatomical images. New spurious features can occur when it is not used properly.

In the absence of any data constraints, the maximization of the Shannon entropy measure, H_S, results in all the pixel values $\rho(l\Delta x)$ being equal to e^{-1}. This fact seems to suggest that the maximum entropy principle in this case selects the most uniform (or the least structural information) image consistent with the data and, therefore, noise and Gibbs ringing can be removed. This point has been demonstrated by examples from applications of these methods to NMR spectroscopic data analysis and radio astronomy image reconstructions. But recent results indicate that this property is not generally true. In fact, it is shown (61) that in some particular cases, the maximum entropy reconstruction is no more than a scaled version of the conventional Fourier reconstruction with a special nonlinear scaling scheme. In (46), it is also demonstrated that maximizing H_S does not guarantee that the smoothest reconstruction will be produced.

Reconstructions from two data sets are presented in Figs. 12 and 13 to illustrate the typical behaviour of MEM for MR image reconstruction. For a comparison, the images

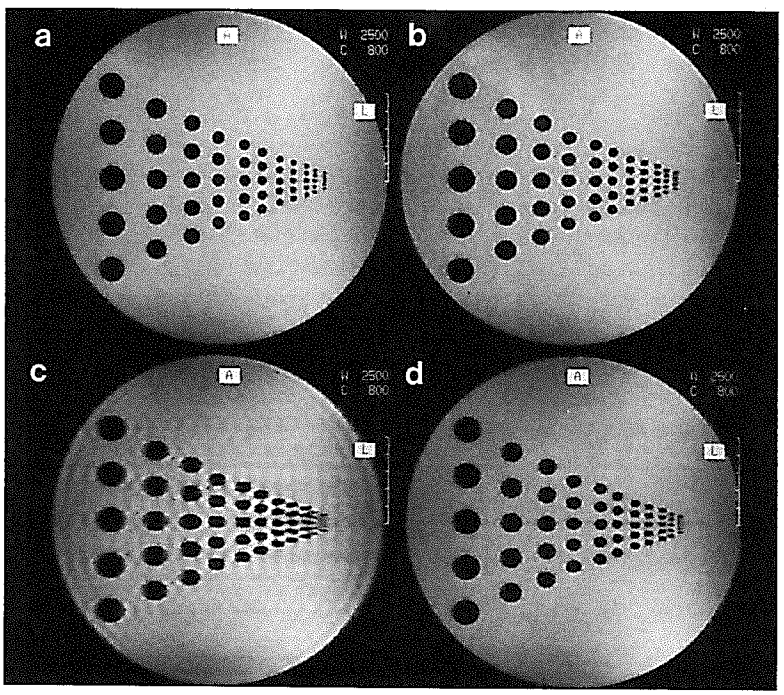

Figure 12. Reconstructions of a phantom image. The vertical direction is reconstructed using the conventional Fourier method with 256 symmetric data points and the horizontal direction is reconstructed using: (a–b) the Fourier method with 256 and 128 data points, respectively; and (c–d) MEM with 96 and 128 data points, respectively. Note the ringing artifacts with the MEM reconstructions, especially in (c). The computational algorithm in (46) was used to produce the MEM images.

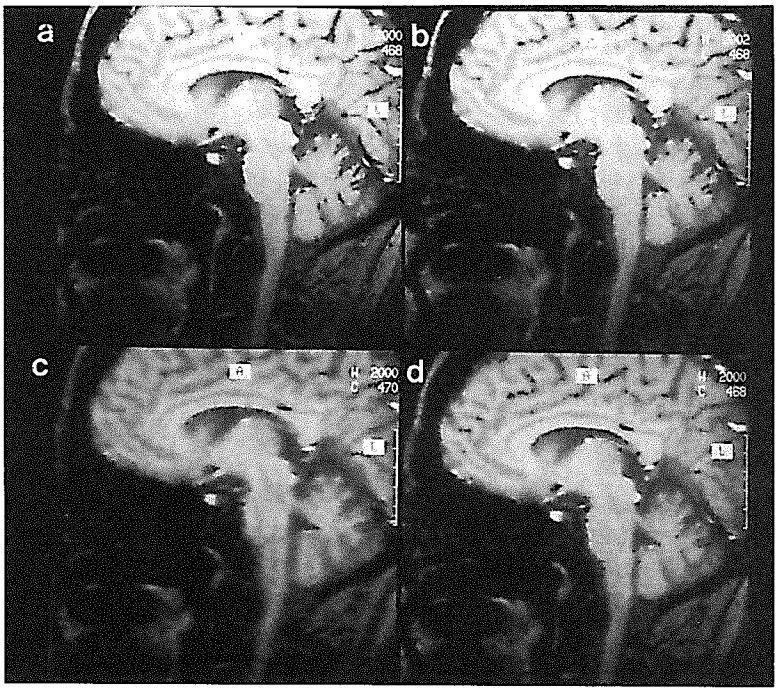

Figure 13. The same as Fig. 12, except for a sagittal head data set.

in Figs. 12a–b and Figs. 13a–b were reconstructed using the standard Fourier method with 256 and 128 symmetric data points, respectively, and the images in Figs. 12c–d and Figs. 13c–d were reconstructed using MEM with 96 and 128 symmetric data points, respectively, by maximizing the Shannon entropy measure H_S. As can be seen, MEM performs poorly in reducing the truncation artifacts in both cases. The Gibbs ringing in the 96 point MEM image of the phantom shown in Fig. 12c is, in fact, increased as compared to that in the 96 point Fourier image in Fig. 4c, suggesting a possible convergence problem with the reconstruction algorithm used (46). Based on our observations of other reconstructions, we concluded that MEM, as it is implemented in (46), is not as useful as other much less computation-intensive methods to be reviewed in this paper for truncated-Fourier imaging. However, we should emphasize that different interpretations and implementations of the maximum entropy principle, tailored to satisfy different image properties through the use of a different Lagrangian formula, might prove useful for other MRI applications (169).

Linear Predictive Extrapolation

Theory. Linear predictive (LP) extrapolation methods are based on the fact that support-limited signals are approximately linear-predictable to a certain order. If enough data points are collected, they will permit the definition of such a linear predictor, which can then take the measured data as input to predict the unmeasured data. To reduce the order of the linear predictor required to model the measured data, it has been found (80,159) to be more effective to apply the modeling to the high-pass filtered data. This is motivated

by the fact that many image functions can be reasonably represented by a series of box car functions (80,82); the derivative of each of them contributes two Dirac δ-functions located at the two edge discontinuities. Accordingly, the Fourier signal is a summation of two complex sinusoids and is exactly second-order linear predictable in the sense that every uniformly sampled data point of the signal can be expressed as a *linear combination* of the previous (or next) two data points with a *fixed* set of weighting coefficients (linear predictor parameters). Based on the differentiation property of the Fourier transform, the original signal from the box car function filtered by $j2\pi k$ will also be second-order linear predictable and, therefore, four data points are sufficient to define the corresponding linear predictor parameters.

In general, the filtered signal satisfies the following linear prediction equation

$$\bar{s}(n\Delta k) = -\sum_{l=1}^{p} h_l \bar{s}[(n-l)\Delta k] + e_n \tag{52}$$

where $\bar{s}(n\Delta k) = i2\pi n\Delta k s(n\Delta k)$, p is the linear predictor order, and e_n is the fitting error. This set of linear-predictive equations is exact ($e_n = 0$) when the image function $\rho(x)$ can be described by a summation of box-car functions (134) (or its derivative is a summation of δ-functions; see the section on the LPA Model for more details). In this case, the h_l can be exactly determined by solving the corresponding linear system if p is less than $N_{data}/2$. In the presence of noise and higher-order (smooth) features, this constraint is less accurate. The linear predictor coefficients, h_l, are now defined in a least squares sense to minimize the total linear prediction errors $\Sigma_n |e_n|^2$. Note that higher-order filters, $(j2\pi k)^L$ for $L > 1$, can reduce the linear prediction errors from higher-order features but will amplify noise and, therefore, may not be useful for practical applications. After the h_l are known, the unavailable data can be predicted using Eq. (52). In principle, infinite extrapolation is possible but is never used in practice because instability of the extrapolation filter defined by $\{h_l\}$ may be exaggerated, leading to unacceptable reconstruction artifacts, notably, high intensity streaks across the image.

Algorithm. Image reconstruction using LP extrapolation consists of the following steps.

1. High-pass filter the measured signal $s(n\Delta k)$ to convert it to a more "linear-predictable" form

$$\bar{s}(k) = i2\pi k s(k). \tag{53}$$

   Higher-order filters can also be used but at the expense of noise amplification.
2. Fit the filtered data with a linear predictor of an appropriate order p. For symmetrically sampled data, the two sides of the echo need to be processed separately. A practical value for p is ⅓ to ½ of the length of each side of the data. The larger the p value the better the spatial resolution, but less stable the extrapolation will be. After the predictor order is chosen, the predictor coefficients $\{h_l, l = 1, 2, \cdots, p\}$ can be calculated using various computational procedures detailed in the appendix.
3. Extrapolate the filtered data using the linear predictor Eq. (52).
4. Multiply the extrapolated signal by $1/(i2\pi n\Delta k)$ to low-pass filter it before appending it to the original signal.
5. Apply the conventional Fourier reconstruction method to the merged data to get the required image.

Characteristics. The properties of the LP extrapolation methods for image reconstruction from finite-Fourier data are summarized as follows.

- The linear predictive constraint is a very powerful one. It can give superresolution reconstruction when it is appropriate for the underlying signals.
- The reconstruction procedure is efficient when the covariance matrix method is used for solving the linear predictor coefficients.
- However, this method is sensitive to noise and the selection of the linear predictor order p. Selection of an optimal linear predictor order is often nontrivial and an unstable linear predictor may result when the order is overdetermined. This is especially true when the covariance matrix method (see the Appendix) is used for solving the LP parameters h_l. Note that the LP extrapolation is equivalent to an all-pole modeling. This can be clearly seen from the equivalent spatial domain representation of Eq. (52) as

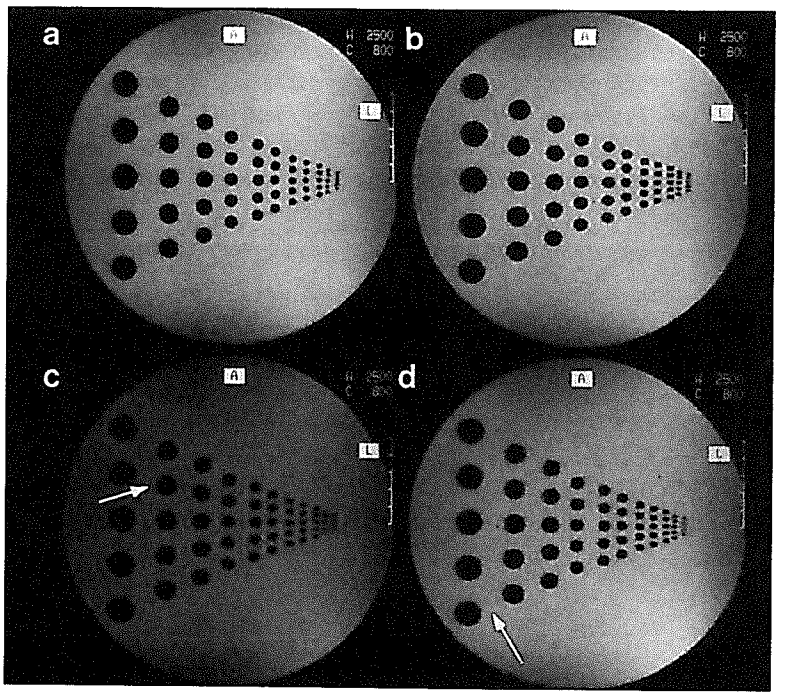

Figure 14. Reconstructions of a phantom image. The vertical direction is reconstructed using the conventional Fourier method with 256 symmetric data points and the horizontal direction is reconstructed using: (a–b) the Fourier method with 256 and 128 data points, respectively; and (c–d) the LP extrapolation method with 96 and 128 data points, respectively. The LP reconstruction is sensitive to the signal-to-noise levels and the number of structures in the image. The 96 point image (c) has trouble removing all the ringing and has many noise spikes (which have been somewhat smoothed in the photographic process). The 128 point image (d) picks up the high resolution features, but also suffers from spike artifacts. Note that the DC reduction in image (c) is due to photography not the reconstruction technique.

$$\tilde{\rho}(x) = \frac{d\rho(x)}{dx} \sim \left. \frac{1}{1 + \sum_{l=1}^{p} h_l z^{-l}} \right|_{z=e^{-i\Delta kx}} . \tag{54}$$

An unstable linear predictor has some poles of the rational function falling on or even outside the unit circle in the z-complex plane, leading to spurious spikes in $\tilde{\rho}(x)$ or high intensity streaks in the final reconstruction (after an effective integration). To overcome this problem, Martin and Tirerdi (159) used the autocorrelation matrix method combined with the Levinson-Durbin algorithm, but at the expense of spatial resolution because of the inherent data windowing with this approach (see Appendix A for detailed discussion). Other SVD-based methods can be used to gain better resolution but with reduced computational efficiency. A more useful LP method, based on the autoregressive moving average model, will be discussed in the next chapter.

Figures 14 and 15 show two sets of reconstructions using the LP extrapolation method in comparison with those from the standard Fourier method. Organized as in Figs. 12, 13, the images in Figs. 14a–b and Figs. 15a–b were reconstructed using the standard FFT method with 256 and 128 symmetric data points, respectively, and the images in Figs. 14c–d and Figs. 15c–d were reconstructed using the LP extrapolation method with 96 and

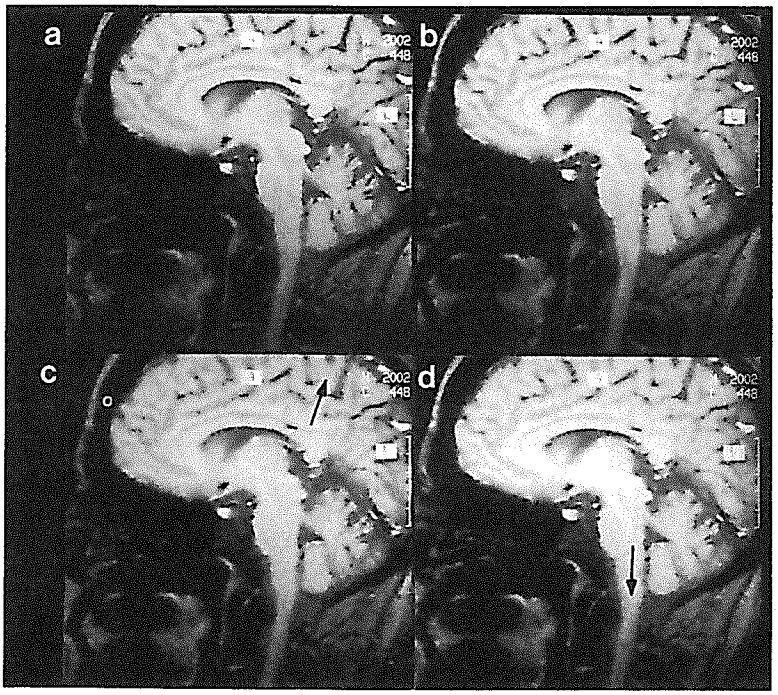

Figure 15. The same as Fig. 14, except for a sagittal head data set. Noise spikes are also visible at different locations in the LP images.

128 symmetric data points, respectively. As can be seen, the LP reconstructions show somewhat reduced Gibbs ringing with improved resolution but suffer noticeable streaking artifacts, particularly in the sagittal head example shown in Figs. 15c–d. These artifacts can be removed using the extended LP methods discussed in the next section.

Sigma Filter Extrapolation

Theory. Image reconstruction using sigma filter (SF) extrapolation is due to Constable and Henkelman (45). This method is based on the following ideas: (a) systematic spurious ringing near sharp edges in an image can be effectively reduced/removed using an edge-preserving filter; and (b) the filtered image, free of spurious Gibbs ringing, can be used to regenerate the missing high spatial-frequency components so that truncation artifacts may be reduced or removed, without the concomitant loss of resolution typically encountered in other filtering schemes for Gibbs ringing reduction.

The key to the success of such an approach lies in the accurate extraction and enhancement of the edge information from the original image containing truncation artifacts and noise. Constable and Henkelman (45) recognized that this can be easily accomplished through the use of a modified sigma filter (132). Like the conventional sliding smoother, a $(2p + 1)$-point sigma filter is defined by a $(2p + 1)$ mask $[w_{-p}(m), \cdots, w_p(m)]$ such that the filtered image is obtained using the following weighted average

$$\tilde{\rho}(m) = \sum_{l=-p}^{p} w_l(m)\rho[(m + l)\Delta x]. \tag{55}$$

For a particular pixel m, the weighting coefficients $w_l(m)$ are defined adaptively based on the local image properties so that sharp edges will not be smoothed. Originally, Lee (132) proposed to choose $w_l(m)$ such that

$$w_l(m) = \frac{\delta_l(m)}{\sum_{l=-p}^{p} \delta_l(m)}, \tag{56}$$

where

$$\delta_l(m) = \begin{cases} 1 & \rho(m\Delta x) - \Delta_\rho \leq \rho[(m + l)\Delta x] \leq \rho(m\Delta x) + \Delta_\rho \\ 0 & \text{otherwise.} \end{cases} \qquad l = [-p,p] \tag{57}$$

Δ_ρ is the edge preserving parameter that determines which pixels are averaged, and it is normally set to be $2\sigma_{imag}$, where σ_{imag} is the standard deviation of the noise in the image. One of the important contributions of the Constable-Henkelman method is the extension of this filter for removal of systematic Gibbs ringing. Specifically, they proposed to (a) use $\Delta_\rho = 2\sigma_{imag} + 0.10\rho(m\Delta x)$ so that the Gibbs overshoots will fall within this threshold (see the *Fourier Reconstruction Method* section for theoretical justifications); and (b) set $p = N_{imag}/N_{data}$ (N_{imag} is the required size of the extrapolated data set) so that the filter width is wide enough to contain the first lobe of the convolution kernel associated with the Fourier transform of the truncated data.

Algorithm. The SF extrapolation algorithm developed by Constable and Henkelman (45) is given below.

1. Zero-fill the measured data to the required extrapolation size and apply the FFT to them to get a zero-padded Fourier reconstruction, $\hat{\rho}_0(m\Delta x)$.
2. Filter the magnitude image $\hat{\rho}_1(m\Delta x) = |\hat{\rho}_0(m\Delta x)|$ using the modified sigma filter described in the previous section to get $\hat{\rho}_2(m\Delta x)$.
3. Edge-enhance $\hat{\rho}_2(m\Delta x)$ by the exponential transformation

$$\hat{\rho}_3(m\Delta x) = [\hat{\rho}_2(m\Delta x)]^{\alpha} \tag{58}$$

   where α is set to 1.5, and rescale $\hat{\rho}_3(m\Delta x)$ so that it has the same dynamic range as $\hat{\rho}_2(m\Delta x)$.
4. Fourier transform $\hat{\rho}_3(m\Delta x)$ back to get a new data set $\tilde{s}(n\Delta k)$ and extrapolate the original data $s(n\Delta k)$ with $\tilde{s}(n\Delta k)$. The linear merging filter described in Eq. (33) is applied to the junctions of the original and new data sets to remove possible discontinuities.
5. Inverse Fourier transform the extrapolated data set to get the final image $\hat{\rho}(m\Delta x)$.

Characteristics. The properties of the SF extrapolation methods for image reconstruction from finite-Fourier data are summarized as follows.

- A distinct advantage of the SF extrapolation method over support-limited extrapolation and linear predictive extrapolation methods is its stability. This method is not sensitive to noise and can achieve simultaneous suppression of Gibbs ringing and noise. It is also computationally very efficient for small filter masks, and can be easily extended to higher-dimensional cases.
- The original Constable-Henkelman SF extrapolation method assumes that the image has negligible phase variations or that reliable phase corrections can be applied to remove the phase before the processing. This assumption may be violated for data sets with large phase variations, as is sometimes the case with gradient-echo experiments, which significantly limits its practical utility. Some remedies have been proposed to overcome this problem. One approach is to use complex sigma filtering in step 2, proposed by Amartur and Haacke (4). In this way, the phase inconsistency (discontinuities) between the measured data $s(n\Delta k)$ and $\tilde{s}(n\Delta k)$ in step 4 can be minimized. A more useful phase handling method suggested by Liang et al. (5) keeps the phase information obtained in step 1, $\varphi_0(m\Delta x)$, and reintroduces it to the filtered magnitude image $\hat{\rho}_3(m\Delta x)$ in step 4 to create $\tilde{s}(n\Delta k)$. This method proves to be nearly as effective as the complex sigma filtering approach, but with a computational time saving of a factor of 2.
- Another subtle point about the SF extrapolation method is that it can give only partial suppression of the Gibbs ringing in a single iteration of the algorithm. This is because the edge information extracted from the Fourier image $\hat{\rho}_0(m\Delta x)$ at step 1 is limited by both the Fourier pixel size and blurring due to convolution with the sinc function. For example, for a step discontinuity, what can be extracted from the Fourier image is a ramp edge going across the step discontinuity. Extrapolation using this ramp edge information can only lead to a reduction of the Gibbs ringing amplitude around this step discontinuity by approximately half. In this case, repeating

steps 2 to 5 iteratively with a relaxed filter parameter Δ_ρ in each iteration as proposed by Amartur and Haacke (4) should be useful.

- Since the SF extrapolation methods extract edge information from the Fourier images, no resolution enhancement with these methods is possible. Instead, edge thickening and loss of subtle image features may occur if the filter parameters are not chosen properly. For this reason, the method is not directly useful for processing asymmetric Fourier data unless the phase constraints are also included for data extrapolation.

Figure 16 shows a set of reconstructions using the original Constable-Henkelman SF extrapolation method and the modified method by Amartur and Haacke in comparison with those from the standard Fourier method. The images in Figs. 16a–b were reconstructed using the standard Fourier method with 256 and 128 data points, respectively. The SF image in Fig. 16c was reconstructed from 128 symmetric data points using the original Constable-Henkelman SF method. As can be seen, this method fails to completely remove the Gibbs ringing because of the large phase variations present in this data set obtained using a gradient-echo sequence (see the phase image in Fig. A1). For

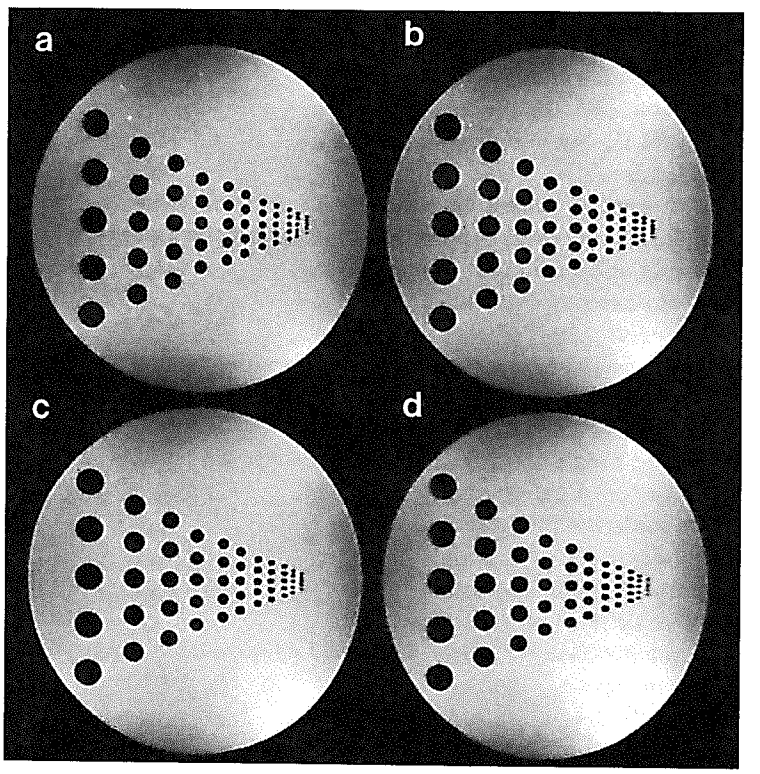

Figure 16. Reconstructions of a phantom image. The vertical direction is reconstructed using the conventional Fourier method with 256 symmetric data points and the horizontal direction is reconstructed using: (a) the FFT with 256 data points; (b) the FFT with 128 data points; (c) the Constable-Henkelman (CH) SF method with 128 data points; and (d) the Amartur-Haacke (AH) SF method with 128 data points.

spin-echo data sets, this may not be a problem. Application of the modified SF extrapolation (4) technique with 3 iterations to the same data sets gave the reconstruction in Fig. 16d. This image shows significantly reduced Gibbs ringing as compared with the 128 point Fourier reconstruction in Fig. 16b, but no resolution improvement, since sigma-filtering *per se* does not lead to better image edge definitions.

Summary

In this section, we have discussed some of the nonparametric methods for finite-Fourier imaging. These methods have in common that they all express the image function in terms of a truncated Fourier series but with a larger number of terms than the number of data points available. The unavailable high-frequency components in the series are recovered using *a priori* constraints. How accurately the missing high spatial frequency data are recovered depends on the constraints themselves and the accuracy of these extrapolated data determines the quality of the final reconstruction. A summary of the salient features of the methods discussed is given below.

1. **Support-limited extrapolation method**
   - *Advantages:*
     — possible resolution improvement;
     — computationally efficient;
     — useful for low-frequency restoration.
   - *Disadvantages:*
     — not generally useful for removal of truncation artifacts in MRI applications because either the support-bound is difficult to determine or it contains no useful information for extrapolation.
2. **Maximum entropy method**
   - *Advantages:*
     — suppression of background noise and spurious image features;
   - *Disadvantages:*
     — computationally very expensive;
     — cannot handle spurious features superimposed on true image features unless the constrained entropy measure is used.
3. **LP extrapolation method**
   - *Advantages:*
     — reduced Gibbs ringing;
     — possible resolution enhancement;
     — computationally efficient when the Levinson-Durbin algorithm is used.
   - *Disadvantages:*
     — instability;
     — new image artifacts, such as spurious streaks;
     — not useful for asymmetric data.
4. **SF extrapolation method**
   - *Advantages:*
     — suppression of Gibbs ringing;
     — possible SNR improvement;
     — efficient and flexible for extension to higher-dimensions;

- *Disadvantages:*
  — no resolution improvement, or resolution may be lost;
  — degraded performance for asymmetric data sets.

PARAMETRIC MODELING METHODS

Introduction

In the previous section, we have discussed several nonparametric methods in which the image function is described in terms of a truncated Fourier series. This nonparametric description has the advantage of being straightforward and, in some cases, computationally efficient. In other cases, incorporating *a priori* constraints into the reconstruction process may become a burden. For example, if an image function $\rho(x)$ is known to be a δ-function, i.e., $\rho(x) = A\delta(x - x_0)$, with an unknown amplitude A and location x_0 then, in principle, two parameters are sufficient to characterize the image. Correspondingly, two data points are supposedly enough to determine the unknowns. However, with the Fourier series, we need to recover the entire series, i.e., an infinite number of data points, to reconstruct the δ-function; a great inefficiency. In this section, we will discuss methods using parametric descriptions of image functions (or the corresponding Fourier signals). While the performance of different models may differ drastically, they share some common properties. First, if a model is appropriate and the number of model parameters is less than the number of data points available, then the parameters can often be determined exactly (in the least squared error sense) and thus the data truncation artifacts can be eliminated. This removes both the Gibbs ringing and the resolution loss. Hence, parametric methods can recover the whole Fourier series without explicitly extrapolating the known data to the infinite frequency range. Another desirable advantage of parametric methods is noise removal. Particularly, if the number of parameters characterizing the model is less than the number of data points, a least squares parameter estimation procedure can effectively "smooth" out some of the measurement random noise. In this section, we will first briefly discuss the general principle and methodology for parametric image reconstruction and then review some of the available models for MRI applications. Other popular image models developed for image recognition and feature extraction are outside the scope of this review and excluded from the discussion.

Principle of parametric modeling. Although parametric modeling has been widely used in various areas of data processing, its application in imaging has been limited because of the complexity of image behavior modeling. In theory, one can parametrize an image function or the corresponding signal in as many ways as one's imagination allows. In practice, the choices are limited by the physical principles (or imaging equation) governing the data collection process and additional *a priori* constraints. When these *a priori* constraints are valid, they often yield a much better model (in terms of efficiency) than that derived from the generic imaging equation. On the other hand, if they are not properly used, the resulting model may bias our interpretation of the finite experimental data, thereby leading to undesirable new image artifacts. The general principles governing the selection of an optimal model are the following (not necessarily exclusive):

- *Sufficiency:* the model can accurately represent the class of image functions of interest;

- *Efficiency/Parsimony:* the model can characterize the image functions with the fewest parameters;
- *Robustness:* the model should be stable with respect to perturbations; including random noise and systematic modeling offsets, of reasonable levels;
- *Computability:* the model parameters can be found efficiently when the model is fitted to the measured data;

The model sufficiency requirement can be satisfied by almost any model if an unlimited number of parameters is used, as in the case of a Fourier series. But, in practice, we are given only a finite number of degrees of freedom in setting up a model, since there is only a finite number of data points available to which the model can be fitted. On the other hand, from the parsimony principle, we always want to have a model with the fewest parameters possible. In this case, the best we can ask for from such a model is that any image function be as accurately approximated as possible. Therefore, the challenge of parametric modeling lies in achieving an optimal compromise between model sufficiency and efficiency (not necessarily computational efficiency). In this regard, *a priori* knowledge of a particular problem as well as the users' ingenuity is most important.

Robustness of a parametric model is important since we always have to deal with measurement random noise and systematic modeling error. Ideally, the model should be stable so that the model parameters are insensitive to perturbations in the assumed working conditions. Computational complexity of a model is also of great importance since *a priori* constraints often require highly nonlinear methods. It is desirable that such a model be computationally as simple as possible so that the parameters can be found efficiently with the available computing technology.

It is possible that no single model is optimal under all these criteria. For example, a common theme in the modeling techniques is to treat the image as a series of rectangular cross sections, as is done with the Haacke technique (82). For certain images, this will result in an efficient model, since sharp edges can be easily described by such a model. However, for an image dominated by higher-order features, such as ramps, this model will become inefficient because a large series of rectangular functions is necessary to represent these features. This comment is true for other modeling techniques when the image does not contain only the stated model characteristics. However, the criteria above do provide some guidelines for establishing an optimal one under some reasonable tradeoffs.

Parametric image reconstruction. The parametric approach to image reconstruction involves three steps. First, an appropriate parametric model is selected. If the model is an accurate representation of the data, an improved reconstruction can be obtained with reduced truncation artifacts; otherwise, biased reconstructions may result. Second, an estimate of the model parameters is made by fitting the model to the available data, which often demands the solution of an optimization problem. If a model is accurate but its coefficients are poorly estimated, a poor reconstruction may also result. It is, therefore, not surprising that there exists a variety of algorithms for a single model for optimal parameter estimation under different conditions. After the model parameters are found, an image is finally generated either directly from the model coefficients, as shown in Fig. 17 (path 1) or from the extrapolated data, as is done in nonparametric methods discussed in the previous section, (path 2 in Fig. 17). The accuracy of the extrapolated data totally

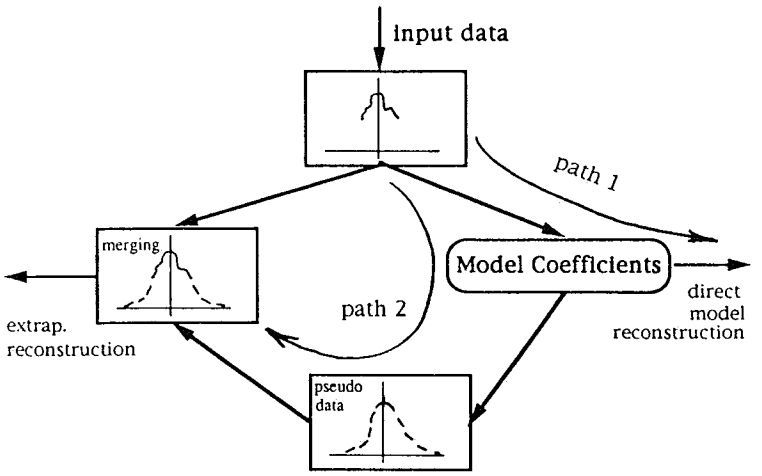

Figure 17. Block diagram of parametric reconstruction methods.

depends on the model itself, which determines the quality of the final reconstruction. This approach sometimes does offer some advantage over the direct model reconstruction, since data consistency can always be maintained (i.e., the final reconstruction is always consistent with the measured data) even if systematic modeling offsets exist. In this way, new reconstruction artifacts can sometimes be avoided if the model constraint is not appropriate for a particular data set. Interestingly, it has been observed that by data extrapolation, some spurious image detail associated with the high frequency noise can be reintroduced to the image to improve its visual quality (66,221). In the following, we will discuss three leading models: The ARMA (autoregressive moving average) model, the LPA (localized polynomial approximation) model, and the GS (generalized series) model for MRI image reconstruction. In order to better appreciate their potential, we will distinguish the models themselves from the algorithms currently employed to calculate the model parameters and the final image. For each model, its characteristics are first discussed in the *theory* subsection and parameter estimations are then described in the *algorithm* subsection, followed by a subsection summarizing the main properties of the model and the algorithms.

The ARMA Model

Theory. The autoregressive moving average (ARMA) model originated in time series modeling (257). It has been widely used in high resolution spectral estimation from short data records in seismic data analysis and various other fields. Application of the model to MR image reconstruction was by Smith et al. (215,216). Because the general ARMA modeling theory is well covered in textbooks and review papers (117,157,186), this section provides only a brief discussion of the necessary concepts for understanding its application to MRI.

In general, a (p,q)-order ARMA signal (or data sequence), $s_{\mathrm{ARMA}}(n\Delta k)$, satisfies the following recursive equation

$$s_{\mathrm{ARMA}}(n\Delta k) = -\sum_{l=1}^{p} a_l s_{\mathrm{ARMA}}[(n-l)\Delta k] + \sum_{l=0}^{q} b_l e[(n-l)\Delta k] \tag{59}$$

where in the statistical literature the a_l are called the autoregressive (AR) coefficients, the b_l the moving average (MA) coefficients, and $e_l = e(l\Delta k)$ are the excitation pulses (random or deterministic). It can be seen that Eq. (59) is a more general form of the linear prediction equation previously discussed in *Linear Predictive Extrapolation* subsection. This model should therefore provide a better framework for linear predictive extrapolation.

In signal processing terms, $s_{\mathrm{ARMA}}(n\Delta k)$ can be viewed as the input generated from a digital filter with transfer function $H(z) = B_q(z)/A_p(z)$ excited by a train of output pulses, e_l, where $A_p(z) = 1 + \Sigma_{l=1}^{p} a_l z^{-l}$ and $B_q(z) = \Sigma_{l=0}^{q} b_l z^{-l}$. Smith et al. (214) treated the MR signal as such a filter output. Either the original signal $s(n\Delta k)$ or the high-pass filtered signal, $i2\pi\Delta k s(n\Delta k)$, can be used. By applying the z-transform to Eq. (59), and using the relationship between the z and Fourier transforms, the ARMA image function (216,219) is given by

$$\begin{aligned}\rho_{\mathrm{ARMA}}(x) &= \mathfrak{I}^{-1}\{s_{\mathrm{ARMA}}(n\Delta k)\} = \Delta k\, E_r(z)\, \frac{B_q(z)}{A_p(z)}\bigg|_{z=e^{-i2\pi\Delta k x}} \\ &= \Delta k \sum_{l=0}^{r} e_l e^{i2\pi\Delta k x} \frac{\sum_{l=0}^{q} b_l e^{i2\pi l\Delta k x}}{1 + \sum_{l=1}^{p} a_l e^{i2\pi l\Delta k x}}\end{aligned} \tag{60}$$

where $E_r(z) = \Sigma_{l=0}^{r} e_l z^{-l}$. Smith et al. have suggested (219) that *a priori* information about the MR signal be incorporated, but currently assume that the excitation sequence e_l is a single pulse, i.e., $e_0 = 1$ and $e_l = 0$, for $l > 0$.

From Eq. (60), one can see that the classical Fourier series model is a strict deterministic MA model [$ARMA(0, N_{data})$; $e_l = \delta(l)$]. As we know, the Fourier series model is optimal for representing smooth changes but is very inefficient for representing spiky image features. On the other hand, an AR model [$1/A_p(z) = ARMA(p, 0)$] is not satisfactory for representing smooth variations but is very effective for representing δ-function like features. A combination of these two capabilities, together with the possibility of introducing *a priori* information via the excitation pulses e_l, should be expected to make the ARMA model more efficient for representing general image features. Especially, because the model has only a finite number of model parameters, the data truncation artifacts can be circumvented if the number of observed data points, N_{data}, adequately exceeds the total number of model parameters ($p + q + 1$), and the algorithm chosen is appropriate for the MR data being modeled. In fact, it can be proved that any continuous function, $\rho(x)$, may be approximated arbitrarily closely by an ARMA model of the form Eq. (60) if the order parameters p and q are selected to be suitably large (40). The widespread use of this model in various fields of data analysis is largely motivated by this property.

Algorithm. Image reconstruction under the ARMA model constraint requires determination of the a_l, b_l and e_l parameters of the model, which are most compatible with measured data $\{s(n\Delta k);\, n \in \mathcal{N}_{data}\}$. The mechanism for measuring this compatibility is normally the following least squares criterion

$$\min_{\{a_l\},\{b_l\},\{e_l\}} \sum_{n\in\mathcal{N}_{data}} \left| s(n\Delta k) + \sum_{l=1}^{p} a_l s_{\mathrm{ARMA}}[(n-l)\Delta k] - \sum_{l=0}^{q} b_l e_{n-l} \right|^2 . \tag{61}$$

If the above least squares solution for the optimal a_l, b_l, e_l parameters is desired, it will almost always necessitate the use of computationally burdensome algorithms with the attendant difficulty of initial parameter value selection and the possibilities of convergence to a local extremum or even nonconvergence. To overcome this problem, all the currently available parameter estimation techniques evaluate the a_l, b_l and e_l separately. In this way, it is possible to provide a linear solution procedure for the a_l parameters. The b_l and e_l parameters can then be determined from the residual data and *a priori* information. One such technique, TERA (Transient Error Reconstruction Algorithm) developed by Smith et al. (216), has found useful applications in MR image reconstruction; the details of it are summarized below. Note that the TERA algorithm has been used to model the original imaging data, $s(n\Delta k)$, and following the work of Haacke et al. (82,135) the high pass filtered signal, $i2\pi n\Delta k s(n\Delta k)$. Except for the final reconstruction of $\rho_{\mathrm{ARMA}}(x)$, the modeling procedure is similar for both signals.

1. *Phase removal:* Although not strictly necessary, phase correction prior to modeling can offer certain advantages, provided the phase estimation can be performed quickly and accurately (162). The dominant linear phase terms can be estimated by checking the echo locations. Removal of higher order phase terms is discussed in (163). By removing these phase variations, the asymmetry in the data, $s(n\Delta k)$, can be reduced so that only one wing of the echo may need modeling, possibly reducing the modeling time by a factor of 2. Removal of the asymmetry is considered important for both the original and the high-pass filtered data (219) to put them into a format more appropriate for modeling by TERA. Even when the asymmetry is not completely removed by phase correction, the asymmetric part of the signal may not need modeling. In addition, phase correction can lead to greater modeling stability, possible reduction of the AR model order or reduction of the modeling error for a given model order. The reduction of the model order is important, as determining the p AR coefficients takes a time of order p^3 (166).

   The disadvantage of phase correction is that it can distort the data prior to its modeling, leading to poorer images, or take longer to perform than the additional modeling it was intended to avoid (162).
2. *Accounting for possible data asymmetries:* For a set of symmetrically sampled data,[12] it is possible to independently model each side of the data, but it is advantageous to split the data into its Hermitian and anti-Hermitian components (216).

[12]Note the difference between data symmetry and sampling asymmetry. For asymmetrically sampled data, the modeling procedure is slightly different; see the next subsection for details.

Assuming that the measured data are $s(n\Delta k)$, $-N_{data}/2 \leq n < N_{data}/2$, the Hermitian $[s_+(n\Delta k)]$ and anti-Hermitian $[s_-(n\Delta k)]$ components can be calculated as

$$\begin{cases} s_+(n\Delta k) = [s(n\Delta k) + s^*(-n\Delta k)]/2, \\ s_-(n\Delta k) = [s(n\Delta k) - s^*(-n\Delta k)]/2, \end{cases} \qquad 0 \leq n < N_{data}/2. \tag{62}$$

The corresponding reconstructions, $\rho_+(n\Delta x)$ and $\rho_-(n\Delta x)$, obtained from $s_+(n\Delta k)$ and $s_-(n\Delta k)$, respectively for $n \geq 0$, can be combined to give

$$\rho(x) = 2\mathrm{Re}\{\rho_+(x)\} + 2i\mathrm{Im}\{\rho_-(x)\} - s(0). \tag{63}$$

Ideally, with no phase error, only $\rho_+(n\Delta x)$ needs to be modeled. In practice $\rho_-(n\Delta x)$ contributes about 10% of the total energy in the final reconstruction (219). It can be modeled with the same AR order p as $\rho_+(n\Delta k)$. However, using a lower order ($\approx p/2$), or no modeling at all, often gives satisfactory results and reduces the computational time by half. Since the modeling procedures for $s_+(n\Delta k)$ and $s_-(n\Delta k)$ are the same, we shall use d_n to represent either for notational simplicity in the following discussions.

3. *Determination of the AR parameters a_l:* Temporarily ignoring the b_l parameters in Eq. (59), the AR coefficients are defined by the linear prediction equation

$$d_n = -\sum_{l=0}^{p} a_l d_{n-l} + \epsilon_n, \qquad n = p+1, \ldots, N_{data}/2 - 1. \tag{64}$$

Determination of the a_l from this set of linear equations is the classical linear prediction problem. One first needs to determine the AR model order p, but finding the correct or the "best" value for p is difficult. Too low an order will lead to loss of resolution and too high an order may cause model instability. Smith et al. (216) suggested choosing p on the order of $N_{data}/6$. Some new fast order estimation techniques discussed in (95,96) promise to provide better estimates of the model order and faster modeling. After p is known, the a_l are determined by minimizing the total forward prediction error i.e., $\min_{\{a_l\}} \Sigma_n |\epsilon_n|^2$. A variety of numerical algorithms are available for solving this problem[13], including the Barrowdale least squares linear prediction algorithm (51), advocated by Smith et al. (216), and other more complex and slower SVD-based algorithms described in the Appendix, which can perform better for signals with lower SNR (as is often the case when modeling is applied to the high-pass filtered data).

The AR filter resulting from these algorithms is often not guaranteed to be stable. Smith et al. (216) proposed increasing the stability of the algorithm by moving the poles away from the unit circle (pole-pulling) using

[13]It is important to emphasize here that some algorithms for estimating AR coefficients, such as the fast Burg lattice algorithm, have an *implicit* combined forward and backward linear prediction formalism, which is not appropriate for this application since such a formalism would incorrectly position the roots of $A_p(z)$, leading to increased rather than decreased truncation artifacts.

$$\tilde{a}_l = a_l \alpha^l; \ \alpha \leq 1.0 \tag{65}$$

as the actual AR filter coefficients. A value of α around 0.99 or lower for very noisy signals is appropriate. Extreme pole-pulling is to be avoided as values of $\alpha \approx 0.95$ will give a TERA reconstruction with truncation artifacts similar to those in a standard Fourier reconstruction. To avoid loss of resolution associated with unnecessary pole-pulling, an efficient automated check of the AR coefficients' stability can be performed after their time consuming calculation (192). Note that the use of $\tilde{a}_l$ would disturb the accuracy of the prediction, but not the final reconstruction, since the MA coefficients b_l are adaptively chosen based on the new AR coefficient values.

4. *Determination of the MA coefficients b_l and the excitation pulses e_l:* Determination of the MA coefficients b_l and excitation pulses e_l is in general difficult. In the TERA algorithm, Smith et al. (214,215,216) assumed that the excitation sequence was a single delta function $e_l = \delta(l)$. In this case, the following simple relationship exists between the MA coefficients and the prediction errors

$$b_n = \epsilon_n. \tag{66}$$

By setting $d_n = 0$ for $n < 0$, the b_n can be calculated as follows:

$$b_n = \begin{cases} d_n + \sum_{l=1}^{p} a_l d_{n-l}, & n \geq p \\ d_n + \sum_{l=1}^{n} a_l d_{n-l}, & n < p. \end{cases} \tag{67}$$

The b_n values for $n < p$, calculated in this way, are typically of the same order of magnitude as the original data. By contrast, the b_n for $n \geq p$ are significantly smaller as they correspond to the difference between the original data and its prediction, thus the name *Transient Error*. Provided that the b_n coefficients decrease more rapidly as a function of n than does the original data, the ARMA with this set of parameters will be successful in reducing the truncation artifacts.

5. *Generation of the ARMA image*[14]: After the a_l and b_l coefficients are available, an image can be explicitly formed from Eq. (60) by calculating each individual pixel value from the model. It is more computationally efficient, however, to evaluate the numerator and the denominator using an FFT algorithm. Zero-padding of the AR and MA coefficients prior to the application of the FFT algorithm is often used to get enough digital resolution to catch the high resolution features embedded in such an image function. Some specialized FFT algorithms, such as those in (217), are useful for improving computational efficiency by removing the computations with the padded zeros. Also, if the modeling is applied to the high-pass filtered data, the resultant ARMA image from Eq. (60) shows only the edges of the underlying image

[14]Note that the terms ARMA image and TERA image are used exchangeably in this review; the former is usually used to emphasize the model and the latter to emphasize the algorithm.

(it is a derivative image) and an integration is required to reconstruct the actual image. Smith et al. (219) observed that the derivative image often needs to be zoomed to a factor of 8–16 prior to integration in order to accurately pick up the super-resolution edges available from modeling the derivative image.

Characteristics. The properties of the ARMA model and the TERA algorithm for image reconstruction from finite-Fourier data are summarized as follows.

- *Flexibility:* The ARMA model is rather flexible. It reduces to the conventional Fourier series model when $p = 0$ and to a strict AR model when $q = 0$. A combination of these two features makes the model more efficient for representing a variety of image features. Particularly, the model has only a finite number of parameters ($p + q + 1$) and can be exactly determined if the number of data points available is large enough. Typically, for symmetrically sampled data, the TERA algorithm requires that $p < N_{data}/6$ and reintroduces all the prediction errors so that effectively $q = N_{data}/2 - 1$ unless some filtering on the MA coefficients is done. Therefore, when the model and the AR algorithm are appropriate, data truncation artifacts can be eliminated with only a finite number of data points. Superresolution is possible, especially for δ-function like image features, since they can be picked up by the denominator of Eq. (60) — the AR part of the model.

  The ARMA model can be applied to both the original imaging data and the high-pass filtered data. In principle, modeling of the high-pass filtered data may be more accurate if the image is dominated by sharp edges, but at the expense of increased computational time. Smith et al. (219) found that direct modeling of the measured imaging data is more efficient and appropriate for a wider class of medical images.

  The TERA algorithm can also be applied to reconstructing partial-Fourier data. Note that in the symmetric data case, the model is applied to the Hermitian and anti-Hermitian components. For asymmetric data, these components cannot be obtained as described above. The approach taken by McGibney and Smith (162,222) for modeling these data is to first obtain a partial-Fourier reconstruction using, for example, the Margosian method described previously; this image is then transformed back into the Fourier domain to create a Hermitian data set, which is processed by the TERA algorithm. This approach may improve resolution over the input partial-Fourier reconstruction where distortion occurs because of data truncation, but its performance is also limited by the quality of the initial partial-Fourier reconstruction.
- *Data consistency and stability:* As the b_l parameters are chosen to produce a match to the MR signal, the modeled signal should never be any worse than the original signal. This data consistency does, to some extent, correct for modeling errors when the underlying image deviates from a deterministic ARMA model. This property led Smith et al. (216) to characterize the TERA algorithm as being able to *"model what it can and then reintroduce everything else."* Another desirable property of this data consistency is that when the AR model order is not properly chosen, unstable poles in $1/A_p(z)$ can be compensated, within the limits of the noise, by additional corresponding zeros introduced into $B_q(z)$ during the determination of the b_l coefficients. These two factors are credited by Smith et al. as the main reasons for the success of the deterministic ARMA model for MRI image reconstruction.

Unfortunately (and understandably), the data consistency alone cannot always guarantee an artifact-free reconstruction when modeling errors are significant. For example, when the TERA algorithm is directly applied to modeling discontinuous image features, such as sharp boundaries, high intensity spikes (hot spots) may occur in the reconstructed image, especially around edges near an object boundary with the background. These spikes may have different phases and appear as undershoots — "cold spots" if the magnitude mode is not used to generate the image from the model. This problem can be alleviated by applying the modeling to the high-pass filtered data, but at the expense of amplified noise. Other methods to overcome this problem include pole-pulling and better model order estimation, as discussed in the *theory* subsection. Remaining instabilities are currently handled by an *ad hoc* technique with some theoretical justification. This approach, called *DFT matching,* is based on the fact that the DFT reconstruction of the truncated data is a reasonable estimate of the final modeled image (223). If the modeling increased the resolution by a factor r, then a given image row from the DFT or modeled images can differ by no more than a certain amplitude determined by the resolution increase, r, and the maximum intensity change in that row. Any modeled pixels that differ by more than this amplitude can be considered as spurious spikes and replaced by the corresponding DFT image pixel value.

- *Computational efficiency:* A major advantage of TERA over the LPA modeling method to be described in the next section is the fact that it can be used directly on the measured phase-encoded data $s(n\Delta k)$. Since these data are less noisy than the high-pass filtered data, it is not necessary to use the slow but extremely stable SVD algorithms to determine the AR coefficients. A time complexity analysis of TERA can be found in (165,166). Modeling using the fast Barrowdale AR algorithm (16,17) is on the order of 5–10 times slower than the FFT reconstruction algorithm; it will be a further order(s) of magnitude slower using a SVD-based algorithm. In some applications, TERA can be applied to modeling localized sections of an image in times equal to the FFT reconstruction times (165,166). Use of fast order estimation algorithms can also further reduce modeling time (95,96).

Figure 18 compares a set of reconstructions using the ARMA model with those obtained from the conventional Fourier method. The images in Figs. 18a–b were reconstructed using the conventional FFT method with 256 and 128 data points, respectively. The ARMA image in Fig. 18c was reconstructed using 128 symmetric data points and an AR model of order $p = 24$. As can be seen, the ARMA reconstruction shows much less ringing than the corresponding 128 point Fourier reconstruction in Fig. 18b. Some resolution improvement over the Fourier reconstruction can also be observed, giving rise to a better definition of the circular structures. Applying the modeling to the corresponding high-pass filtered data set, as described above, gave the image in Fig. 18d. This image is clearly inferior to the one in Fig. 19c in terms of resolution and ringing suppression, suggesting that the ARMA modeling should be applied to the unfiltered data, as Smith et al. originally proposed (216). Unless otherwise noted, all other ARMA images presented in this review are reconstructed using this approach.

Figure 19 compares a set of partial-Fourier reconstructions using the ARMA model with those obtained from the conventional Fourier method. As organized in Fig. 18, the images in Figs. 19a–b were reconstructed using the conventional Fourier method with 256 and 128 symmetric data points, respectively. The ARMA images in Figs. 19c–d were

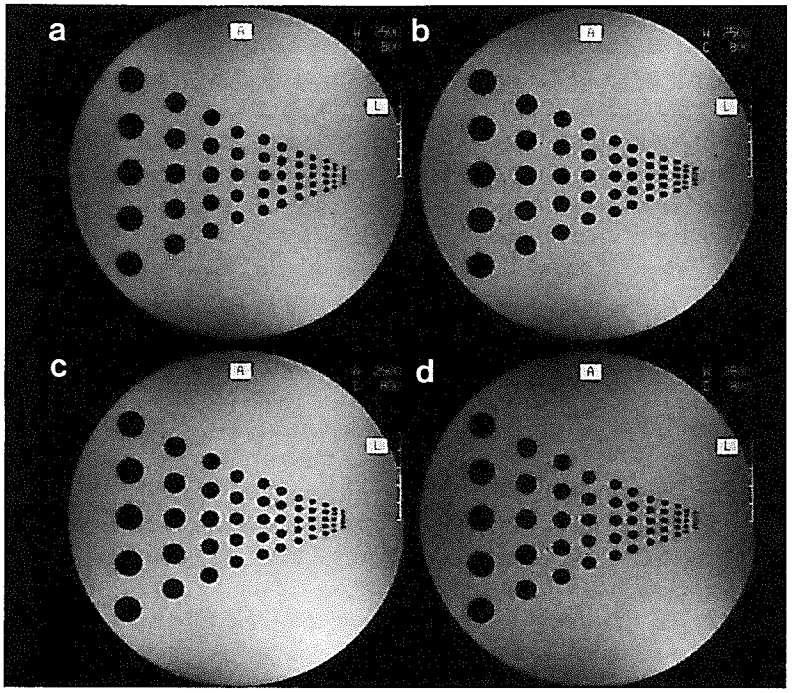

Figure 18. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a–b) the FFT with 256 and 128 symmetric data points, respectively; and (c–d) the TERA algorithm with 128 symmetric data points, without and with high-pass filtering.

reconstructed using 128 asymmetric data points with $n_0 = 16$ and 32 data points before the echo, respectively. The initial input to the TERA algorithm was obtained from a Margosian partial-Fourier reconstruction. As can be seen, the ARMA partial-Fourier reconstructions give much better resolution than the Fourier reconstruction from 128 symmetric data points, but some "phase artifacts" are also obvious in Fig. 19c where the number of data points before the echo is small ($n_0 = 16$). These artifacts are less serious than those in the corresponding Margosian image in Fig. 9. Both ARMA images suffer a noticeable loss in signal-to-noise ratio.

The LPA Model

Theory. While the ARMA model is powerful for characterizing both continuous and δ-function like image features, it is inefficient for representing discontinuous image features, particularly sharp edges. The LPA (localized polynomial approximation) originated by Haacke et al. (82,135) was, to some extent, developed to overcome this problem. This model primarily stems from the following considerations: (a) discontinuities of an image function, $\rho(x)$, are the source of the Gibbs ringing; it is, therefore, crucial for a model to be able to "efficiently" characterize such behavior so that when the model

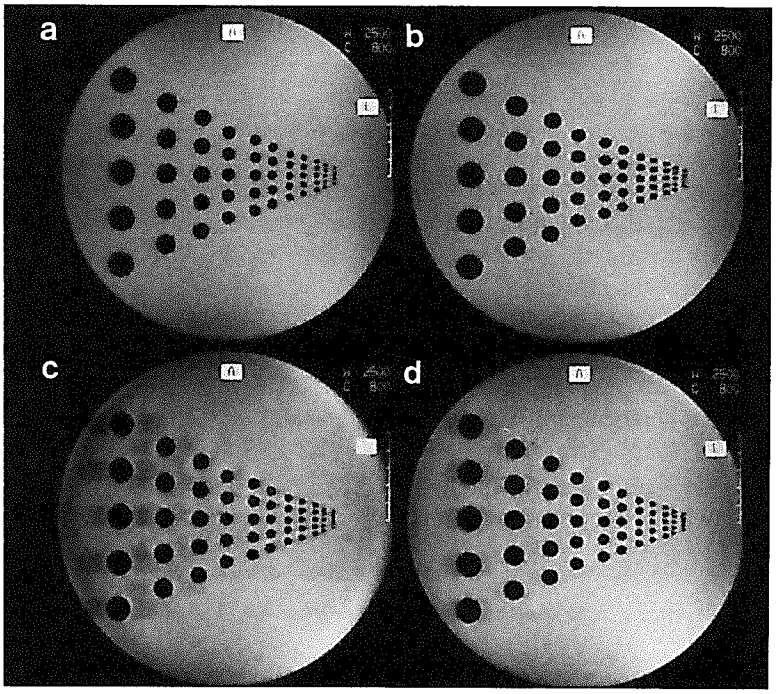

Figure 19. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a–b) the FFT with 256 and 128 symmetric data points, respectively; (c) the TERA algorithm with 128 asymmetric data points (n_0 = 16 points before the echo); and (d) the same as (c) but with n_0 = 32 points before the echo. For both images, the Margosian partial-Fourier reconstruction was used as an initial input. Note the residual phase artifacts.

is fitted to the data, information embedded in the data will be best utilized for the recovery of the true image function; and (b) a piecewise continuous function that is sufficiently smooth in each local region can be accurately represented in a local Taylor series and possibly with a smaller truncation error than with the Fourier series, since such an expansion does not impose periodicity, which can create discontinuities at the ends of the local intervals as the Fourier series expansion does. Mathematically, the model adaptively divides the object support into M local regions, (α_l, α_{l+1}), for $l = 1, 2, \ldots, M$, in which the image function $\rho(x)$ is approximated with a polynomial (truncated Taylor series) of a certain order such that

$$\rho_{\mathrm{LPA}}(x) = \sum_{l=1}^{M} \left[\sum_{r=0}^{P_l} c_{lr} \left(x - \frac{\alpha_l + \alpha_{l+1}}{2} \right)^r \right] w_l(x), \tag{68}$$

where

$$w_l(x) = \begin{cases} 1, & \text{for } \alpha_l < x < \alpha_{l+1} \\ 0, & \text{otherwise.} \end{cases} \tag{69}$$

In the Fourier domain, the model becomes

$$s_{\mathrm{LPA}}(k) = \sum_{l=1}^{M} \sum_{r=0}^{P_l} c_{lr} \Phi_{lr}(k) \tag{70}$$

where the basis functions $\Phi_{lr}(k)$ are given by

$$\Phi_{lr}(k) = \int_{\alpha_l}^{\alpha_{l+1}} \left(x - \frac{\alpha_l + \alpha_{l+1}}{2} \right)^r e^{-i2\pi kx} dx. \tag{71}$$

The independent parameters of an LPA model include

M: *the number of local regions*;
$\{\alpha_l\}$: *the edge locations determining each local region*;
$\{P_l\}$: *the local polynomial orders*; *and*
$\{c_{lr}\}$: *the local polynomial coefficients*.

If the maximum local polynomial order allowed in the model is $P = max\{P_1, \ldots, PM\} = P$, Eq. (68) is called a P^{th}-order LPA model. If $P = 0$, $\rho_{\mathrm{LPA}}(x)$ reduces to a box car model that consists of a series of box car functions with arbitrary edge locations and amplitudes (82). This model, simple and restrictive as it may seem, is a powerful constraint to account for step changes in an image function. One can argue that smooth higher-order changes can also be rebuilt using this model. The underlying principle is similar to what is used in the popular Simpson formula for numerical integration, except that the step sizes and the number of steps (intervals) are not fixed, but adaptively chosen based on the problem. The LPA model, in general, offers the following desirable properties for finite-Fourier imaging.

1. *Capability for modeling piecewise continuous behavior:* The capability to model piecewise continuous behavior of an image function is an immediate result of using localized basis functions (87). Hence, the problem of modeling discontinuities with nonlocal basis functions, such as the complex sinusoids used in the Fourier series, can be easily overcome. Particularly, the LPA model parameterizes the edge locations into the basis functions, enabling the model to pick up discontinuities without being subject to the Fourier pixel-limited resolution. This may be an advantage over the sigma-filtering approach for extracting image edge information discussed in the *Sigma Filter Extrapolation* section.
2. *Localized modeling errors:* Another advantage of using localized basis functions comes from the fact that, when the model is fitted to the data, the image function will be approximated separately in different intervals by the corresponding localized polynomials; modeling errors from one interval may not be propagated to the other intervals through the basis functions. This is an advantage over using the nonlocal basis functions, since it makes it more straightforward to interpret the reconstructions as well as possible artifacts.

3. *Optimized modeling efficiency:* As pointed out earlier given a limited number of data points, it is paramount for a basis function set to be able to model an image function using only a small number of terms. The use of localized basis functions is an important step for achieving this since it is usually very easy to model the local behavior of an image function with a few parameters, provided the local intervals are properly chosen. Particularly, when the image function $\rho(x)$ contains step discontinuities, the convergence rate of the Fourier series will be $1/N$, with N being the number of terms used. However, using the LPA modeling approach, if the local regions are optimally defined, it is possible to achieve a local convergence rate of $1/(1 \times 2 \times \cdots \times P_l)$ with P_l being the order of the local polynomial fitted. Therefore, it is potentially more efficient to use an LPA model to characterize a piecewise continuous image function than to use the Fourier series.

Algorithm. Given a set of sampled points for $s(k_n)$, for $n = 1, 2, \cdots, N_{data}$, one needs to find all the model parameters, including $\{\alpha_l\}$ for the edge locations and $\{c_{lr}\}$ for the polynomial coefficients. As in the ARMA modeling case, $\{\alpha_l\}$ and $\{c_{lr}\}$ should ideally be determined by solving the least squares problem

$$\min_{\{\alpha_l\},\{c_{lr}\}} \sum_{m=1}^{N_{data}} |r_m|^2 = \min_{\{\alpha_l\},\{c_{lr}\}} \sum_{m=1}^{N_{data}} |s(k_m) - s_{\mathrm{LPA}}(k_m)|^2, \tag{72}$$

or in matrix form

$$\min_{\{\alpha_l\},\{c_{lr}\}} \|\boldsymbol{r}\| = \min_{\{\alpha_l\},\{c_{lr}\}} \|\boldsymbol{s} - \boldsymbol{\Phi c}\|, \tag{73}$$

where

$$\boldsymbol{c} = [c_{10}, c_{11}, \ldots, c_{MP_M}]^t$$
$$\boldsymbol{s} = [s(k_1), s(k_2), \ldots, s(k_{N_{data}})]^t$$

and

$$\boldsymbol{\Phi} = \begin{bmatrix} \Phi_{10}(k_1) & \cdots & \Phi_{lr}(k_1) & \Phi_{l(r+1)}(k_1) & \cdots & \Phi_{MP_M}(k_1) \\ \Phi_{10}(k_2) & \cdots & \Phi_{lr}(k_2) & \Phi_{l(r+1)}(k_2) & \cdots & \Phi_{MP_M}(k_2) \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \Phi_{10}(k_{N_{data}}) & \cdots & \Phi_{lr}(k_{N_{data}}) & \Phi_{l(r+1)}(k_{N_{data}}) & \cdots & \Phi_{MP_M}(k_{N_{data}}) \end{bmatrix}, \tag{74}$$

is called the LPA- model matrix, with $\Phi_{lr}(k_n)$ being defined by Eq. (71). Because of the highly nonlinear nature of the problem, existing algorithms reduce the computational requirements by resorting to asymptotic least-squares solutions. Notably, they decouple the determination of the nonlinear parameters $\{\alpha_l\}$ from the linear ones $\{c_{lr}\}$ to make the computational procedure more practical. After the $\{\alpha_l\}$ are known, the $\{c_{lr}\}$ can be easily determined using a standard linear least squares fitting method such that (131)

$$\boldsymbol{c} = (\boldsymbol{\Phi}^H\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^H\boldsymbol{r} = \boldsymbol{\Phi}^{+}\boldsymbol{r}. \tag{75}$$

In the following, two existing algorithms for finding the $\{\alpha_l\}$ are described.

Linear predictive algorithm — The linear predictive algorithm developed by Liang, Haacke, and Thomas (135) is based on the observation that, by taking a $(P + 1)^{th}$-order derivative, a P^{th}-order LPA model can be converted to a generalized δ-function model consisting of a summation of Dirac δ-functions and their derivatives. The Fourier signal generated from such a model becomes linearly predictable. To avoid noise amplification, the first-order derivative is often used. Using this linear predictability, one can determine the nonlinear parameters $\{\alpha_l\}$ by solving a set of linear prediction equations and a polynomial equation. The basic idea of this algorithm is best understood by considering the simplest LPA model consisting of a single box car function

$$\rho_{\mathrm{LPA}}(x) = \begin{cases} A, & \text{for } x_1 < x < x_2 \\ 0, & \text{otherwise.} \end{cases} \tag{76}$$

The first-order derivative model becomes two δ-functions located at x_1 and x_2

$$\frac{d\rho_{\mathrm{LPA}}(x)}{dx} = A\delta(x - x_1) - A\delta(x - x_2), \tag{77}$$

and the corresponding Fourier signal becomes

$$\tilde{s}(n\Delta k) = -i2\pi n\Delta k s(n\Delta k) = Ae^{-i2\pi n\Delta k x_1} - Ae^{-i2\pi n\Delta k x_2}. \tag{78}$$

It is easy to prove that the data sequence $\{\tilde{s}(n\Delta k)\}$ is second-order linear-predictable, i.e.,

$$\tilde{s}(n\Delta k) = -h_1\tilde{s}[(n - 1)\Delta k] - h_2\tilde{s}[(n - 2)\Delta k], \qquad \forall n \tag{79}$$

with

$$\begin{cases} h_1 = -e^{-i2\pi\Delta k x_1} - e^{-i2\pi\Delta k x_2} \\ h_2 = e^{-i2\pi\Delta k(x_1 + x_2)}. \end{cases} \tag{80}$$

The characteristic polynomial of this linear prediction equation is

$$\mathcal{P}(z) = 1 + h_1 z^{-1} + h_2 z^{-2}, \tag{81}$$

whose roots are related to the edge locations x_1 and x_2 simply by

$$z_1 = e^{-i2\pi\Delta k x_1}, \quad \text{and} \quad z_2 = e^{-i2\pi\Delta k x_2}. \tag{82}$$

One may recall that Eq. (79) is the foundation for the linear predictive extrapolation method discussed previously. Here, this LP equation is used as an intermediate step for finding the edge locations x_1 and x_2. Representing z_l as $e^{r_l + i\beta_l}$, instability of $\mathcal{P}(z)$ results if some r_l values are larger than zero (i.e., z_l outside the unit circle of the complex z-plane). This instability, however, will not translate to LPA reconstruction artifacts,

since the r_l are not entered into the edge location calculation. This is an advantage over other LP-based reconstruction algorithms, since the r_l are more sensitive to noise than the β_l. In the following, the LP algorithm is summarized.[15]

1. High-pass filter the measured signal by multiplying it with $i2\pi n\Delta k$ or $(-i2\pi n\Delta k)$:

$$\tilde{s}(n\Delta k) = i2\pi n\Delta k s(n\Delta k). \tag{83}$$

2. Fit the filtered data with an $\tilde{M}$-order linear predictor defined by $\{h_l\}$ such that

$$\tilde{s}(n\Delta k) = \sum_{l=1}^{\tilde{M}} h_l \tilde{s}[(n-l)\Delta k] + e_n \tag{84}$$

where $\tilde{M}$ is often chosen to be $N_{data}/2$ and the h_l can be found using a Levinson or SVD-based algorithm (see the Appendix for more details). Note that, unlike in ARMA modeling, separate modeling of each side of the echos is not necessary for symmetric data.

3. Solve the following characteristic polynomial equation for all its roots $\{z_l,\ l = 1, \cdots, \tilde{M}\}$:

$$\mathcal{P}(z) = 1 + h_1 z^{-1} + \cdots + h_{\tilde{M}} z^{-\tilde{M}} = 0 \tag{85}$$

4. Extract the edge parameters α_l from the roots of $\mathcal{P}(z)$ as:

$$\alpha_l = -\frac{1}{2\pi\Delta k} \operatorname{Im}[\log_e(z_l)]|_{\operatorname{mod}[-\pi,\pi]}, \text{ for } l = 1, 2, \ldots, \tilde{M} \tag{86}$$

5. Sort the calculated edges in an increasing order and remove clustered edges to avoid model instability by lumping together edges if their separations are below a predetermined threshold δ_α, which is chosen to be a fraction of the Fourier pixel size. The number of edges left, $(M + 1)$, determines the number of local intervals fitted. Note that M corresponds to the p parameter in the ARMA model, but the latter is more difficult to choose.
6. Determine the local polynomials orders, P_l, for $l = 1, 2, \ldots, M$, based on the size of the local intervals. It is suggested that for a P^{th}-order model (135)

$$P_l = \min\{P, \max\{0, (\alpha_l - \alpha_{l+1})/\Delta x_f - 1\}\} \tag{87}$$

where Δx_f is the Fourier pixel size, i.e., $1/(N_{data}\Delta k)$ and P is chosen to be 0 to 2.

Variable projection algorithm — The LP estimation of the nonlinear parameters $\{\alpha_l\}$ is statistically biased. When the data signal-to-noise ratio is low, the bias on the estimated model parameters can sometimes be large enough to compromise the quality of the reconstructed result. To decrease the estimation bias, and hence obtain a better fit to the

[15]Note that if the HSVD algorithm described in the Appendix is used for LP fitting, the polynomial equation need not be solved directly and steps 2 and 3 in the algorithm should be modified accordingly.

data, an unbiased nonlinear least squares optimization procedure should be used. Such a procedure is also advantageous because we can now operate on the original data, which does not suffer from the same noise amplification problems as the high-pass filtered data $\bar{s}(n\Delta k)$. To solve the problem formulated in Eq. (72) effectively, a variable projection (VARPRO) technique (77) can be used. This technique is based on the idea that if we replace $\boldsymbol{c}$ by $\boldsymbol{\Phi}^{+}\boldsymbol{s}$, the resultant error function, called the *variable projection functional* $\|\boldsymbol{r}_{VP}\|$, is given by (72,77)

$$\boldsymbol{r}_{VP} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{+}\boldsymbol{s} - \boldsymbol{s} = (\boldsymbol{\Phi}\boldsymbol{\Phi}^{+} - \boldsymbol{I})\boldsymbol{s} \tag{88}$$

and is dependent only on $\boldsymbol{\alpha}$. This significantly reduces the dimension of the parameter space, thereby speeding up the convergence. The optimal $\boldsymbol{\alpha}$ can be determined from this functional by using an iterative search method such that in the pth iteration

$$\boldsymbol{\alpha}_p = \boldsymbol{\alpha}_{p-1} + \Delta\boldsymbol{\alpha}, \tag{89}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{M+1}]^t$. The initial estimate, $\boldsymbol{\alpha}_0$, can be obtained using the above described LP algorithm. The search direction and step size $\Delta\boldsymbol{\alpha}$ can be chosen in each iteration using the modified Gauss-Newton method such that

$$\Delta\boldsymbol{\alpha} = -(\boldsymbol{J}^H\boldsymbol{J} + \mu\boldsymbol{I})^{-1}\boldsymbol{J}^H\boldsymbol{r}_{VP}, \tag{90}$$

where $\boldsymbol{J}$ is the Jacobian matrix defined by

$$\boldsymbol{J} = \left[\frac{\partial \boldsymbol{r}_{VP}}{\partial\alpha_1}, \frac{\partial \boldsymbol{r}_{VP}}{\partial\alpha_2}, \ldots, \frac{\partial \boldsymbol{r}_{VP}}{\partial\alpha_{M+1}}\right], \tag{91}$$

and μ is the regularization parameter usually chosen to make the modified matrix $\boldsymbol{J}^H\boldsymbol{J} + \mu\boldsymbol{I}$ positive definite. Superimposing the realness constraint on $\Delta\boldsymbol{\alpha}$, one further obtains

$$\Delta\boldsymbol{\alpha} = -[\boldsymbol{J}^H\boldsymbol{J} + (\boldsymbol{J}^H\boldsymbol{J})^* + \mu\boldsymbol{I}]^{-1}[\boldsymbol{J}^H\boldsymbol{r}_{VP} + (\boldsymbol{J}^H\boldsymbol{r}_{VP})^*]. \tag{92}$$

Representing $\boldsymbol{\Phi}$ in terms of the QR decomposition, i.e.,

$$\boldsymbol{\Phi} = \boldsymbol{Q}^H\begin{bmatrix}\mathbf{U}\\ \mathbf{0}\end{bmatrix}$$

where $\boldsymbol{Q}$ is an $N_{data} \times N_{data}$ orthonormal matrix and $\boldsymbol{U}$ is an $N_c \times N_c$ upper triangular matrix with N_c being the length of the coefficient vector $\boldsymbol{c}$, one can prove that the $\Delta\boldsymbol{\alpha}$ given above will not be affected by replacing $\boldsymbol{J}$ and $\boldsymbol{r}_{VP}$ with $\boldsymbol{J}_Q = \boldsymbol{Q}\boldsymbol{J}$ and $\boldsymbol{r}_Q = \boldsymbol{Q}\boldsymbol{r}_{VP}$. Therefore, the key step in calculating $\Delta\boldsymbol{\alpha}$ is to find $\boldsymbol{Q}(\partial\boldsymbol{r}_{VP})/(\partial\alpha_l)$, which involves the differentiation of the pseudoinverse matrix $\boldsymbol{\Phi}^{+}$ (72) if one proceeds from Eq. (88). Using a simpler derivation, it is found that (134)

$$\boldsymbol{Q}\frac{\partial \boldsymbol{r}_{VP}}{\partial\alpha_l} = -[\boldsymbol{U}^{-1}\ \mathbf{0}]^H\left(\frac{\partial\boldsymbol{\Phi}}{\partial\alpha_l}\right)^H\boldsymbol{r}_{VP} + \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\boldsymbol{Q}\frac{\partial\boldsymbol{\Phi}}{\partial\alpha_l}\boldsymbol{c}, \tag{93}$$

where $\boldsymbol{I}$ is $(N_{data} - N_c)^{th}$-order identity matrix. $\boldsymbol{Q}(\partial \boldsymbol{r}_{VP})/(\partial \alpha_l)$ can be calculated efficiently from the above formula since the matrix $(\partial \boldsymbol{\Phi})/(\partial \alpha_l)$ is sparse with at most $P_l + P_{l+1} + 2$ non-zero columns. In summary, the VARPRO algorithm for finding $\boldsymbol{\alpha}$ is given as follows:

1. Get an initial estimate for $\boldsymbol{\alpha}$ via, for example, the LP algorithm.
2. Calculate the LPA-model matrix $\boldsymbol{\Phi}$, Eq. (74).
3. Perform a QR decomposition on $\boldsymbol{\Phi}$ to get $\boldsymbol{Q}$ and $\boldsymbol{U}$.
4. Calculate the Jacobian matrix $\boldsymbol{J}$ (actually, $\boldsymbol{J}_Q = \boldsymbol{QJ}$) based on Eqs. (91) and (93).
5. Calculate the stepsize $\Delta\boldsymbol{\alpha}$, based on Eq. (92) but replacing $\boldsymbol{J}$ and $\boldsymbol{r}_{VP}$ with $\boldsymbol{J}_Q$ and $\boldsymbol{r}_Q$, respectively.
6. Update $\boldsymbol{\alpha}$ with $\boldsymbol{\alpha} + \Delta\boldsymbol{\alpha}$ if $\|\boldsymbol{r}_{VP}\|$ is decreased and go back to step 2 if the maximum number of iterations is not reached, otherwise accept the current $\boldsymbol{\alpha}$ as the final output.

Characteristics. The properties of the LPA model for image reconstruction from finite-Fourier data are summarized as follows.

- *Flexibility:* The LPA model constraint is a rather powerful one. Its usefulness lies in its unique capability to represent sharp image edges, and it may yield superresolution reconstruction for image features with the model characteristics described. Particularly, dominant image edges can often be accurately determined with either the LP or the VARPRO algorithms and spurious ringing from them can be suppressed or even eliminated. In practice, the zeroth-order LPA model is often used for handling step discontinuities. This model can pick up sharp, distinct image features like fat in T_1-weighted images, or blood vessels, and should prove useful for angiographic imaging.

  The LPA reconstruction algorithms described above can be applied directly to either symmetric or asymmetric data sets and should be useful for both spin-echo and gradient-echo imaging. When they are applied to asymmetric data, combined forward and backward LP can be used for improved accuracy for the estimates of the edge locations α_l. Also, as in ARMA modeling, it is often helpful to obtain a partial-Fourier image first by using any of the methods described in the *Phase-Constrained Reconstruction* section, and then to estimate the edge locations from the Hermitian data set generated from this image (in this case, the combined forward and backward LP is not necessary). To avoid distortions, the amplitude parameters should be calculated by fitting the model with the estimated edge parameters to the original data. After all the parameters are available, the model can be used to generate the missing data, followed by the conventional Fourier reconstruction.
- *Data consistency and stability:* In practice, the LPA model reconstruction may not be data consistent. This is especially true when the zeroth-order LPA model is used for representing higher-order image features. This data inconsistency manifests itself as streaking artifacts in the direct model reconstruction. To enforce the data consistency constraints, it was suggested (28) that iterative application of the model to the residual data be used to successively pick up image features at different levels without sacrificing model stability, but at the expense of the already low computational efficiency. A better approach is to use the LPA model for data extrapolation so that data consistency is maintained in the final reconstruction (82). To avoid

possible discontinuities at the transitions between the extrapolated data and the measured data, the linear merging filter described in Eq. (33) should be applied.

The LPA model is usually stable or can easily be stabilized by eliminating very narrow boxes (probably at the expense of resolution). Parameter estimation using the LP algorithm is rather robust and does not have the instability problem of other LP-based reconstruction techniques. This is because the LPA model reconstruction is not directly related to the stability of the LP characteristic polynomial. As discussed in the previous section, all the roots (stable or unstable) of $\mathcal{P}(z)$ are projected onto the unit-circle of the z-plane for determination of the edge locations, and closely clustered edges are eliminated to avoid noise spikes. In addition, the high-pass filtering effect (step 1 of the LP algorithm) can be well-balanced out by a low-pass "numerical filtering" effect introduced by using a truncated-SVD-procedure for finding the LP coefficients. This is especially true with symmetric data, since the two filtering effects will exactly cancel out in this case. However, when the data SNR is low, subtle image edges will be buried in the noise because of the increased noise variance and will be perturbed or even removed by the SVD numerical filtering. In this case, the VARPRO algorithm is useful for refining the estimation of model parameters.

- *Computational efficiency:* LPA model-based reconstruction is relatively inefficient with the current form of parameter estimation algorithms, especially when the SVD

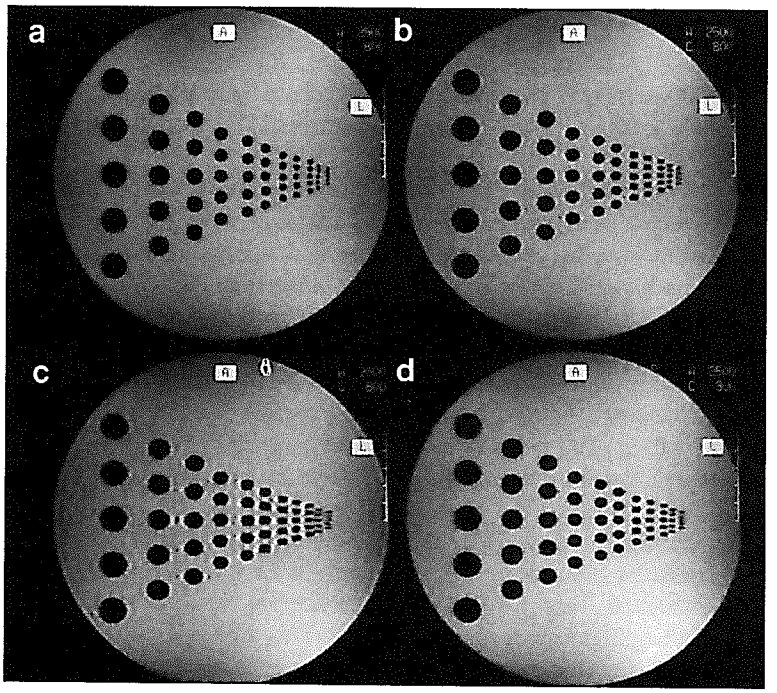

Figure 20. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a–b) the FFT with 256 and 128 symmetric data points, respectively; and (c–d) the LPA model with 128 symmetric data points using the LP and VARPRO parameter estimation algorithms, respectively.

is used for finding the LP coefficients. An LPA model reconstruction using the existing algorithms is about two orders of magnitude slower than a Fourier reconstruction, which is probably the major disadvantage of this reconstruction algorithm.

Figure 20 compares a set of reconstructions obtained using the LPA model with those from the conventional Fourier method. The images in Figs. 20a–b were reconstructed using the standard Fourier method with 256 and 128 data points, respectively. The images in Figs. 20c–d were reconstructed using the zeroth-order LPA model with 128 symmetric data points, using the LP algorithm and the VARPRO algorithm, respectively. As can be seen, both images show excellent suppression of the Gibbs ringing. This is somewhat understandable since the sharp edges in the phantom are very well represented by the LPA model. Some degree of resolution improvement can also be observed, leading to better edge definitions for the circular holes. Image (d) is slightly better than image (c) because of the extra optimization step in refining the model parameters. Both images have good signal-to-noise ratio and no spurious reconstruction artifacts can be observed.

Figure 21 compares a set of partial-Fourier reconstructions obtained using the LPA model with those from the conventional Fourier method. As in Fig. 20, the images in Figs. 21a–b were reconstructed using the conventional FFT method with 256 and 128 symmetric data points, respectively. The LPA images in Figs. 21c–d were reconstructed using 128 asymmetric data points with $n_0 = 16$ and 32 data points before the echo,

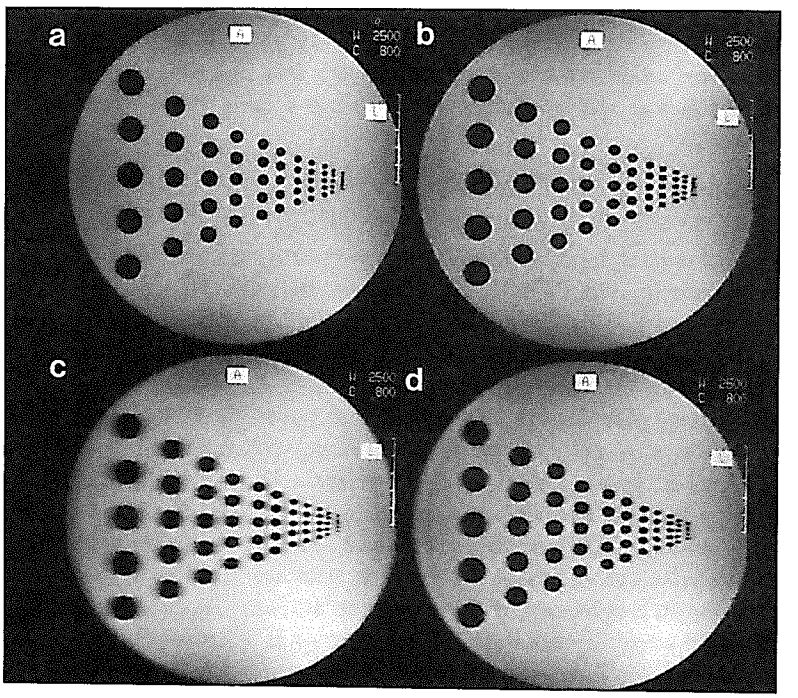

Figure 21. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a–b) the FFT with 256 and 128 symmetric data points, respectively; (c) the LPA model with 128 asymmetric data points ($n_0 = 16$ points before the echo); and (d) the same as (c) but with $n_0 = 32$ points before the echo.

respectively. As can be seen, both LPA partial-Fourier reconstructions give better resolution than the Fourier reconstruction from 128 symmetric data points, without noticeable systematic artifacts or loss of signal-to-noise ratio that often occur with other partial-Fourier reconstruction techniques.

The Generalized Series Model

Theory. The generalized series (GS) model was developed by Liang and Lauterbur (136) in an attempt to establish a general mathematical framework for handling *a priori* constraints in the most data-consistent way. In this model, an image function is represented as

$$\rho_{GS}(x) = \sum_l a_l \varphi_l(\boldsymbol{\theta}, x), \tag{94}$$

where $\varphi_l(\boldsymbol{\theta},x)$ are parametrized basis functions with parameter vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]$ adaptively chosen for optimal modeling in a particular application; a_l are the series coefficients chosen to match the measured data. Of particular interest is the class of basis functions proposed by Liang and Lauterbur (136), which is formed from the family of weighted complex sinusoids as

$$\varphi_l(\boldsymbol{\theta},x) = \mathcal{C}(\boldsymbol{\theta},x)e^{i2\pi l\Delta kx}, \tag{95}$$

where $\mathcal{C}(\boldsymbol{\theta},x)$ is a constraint function, which can take a variety of forms for incorporating available *a priori* information. For a stable model, it is suggested (145) that $\mathcal{C}(\boldsymbol{\theta},x)$ be a non-negative function. Based on this set of basis functions, the GS model becomes

$$\rho_{GS}(x) = \mathcal{C}(\boldsymbol{\theta},x) \sum_{l\in\mathcal{N}_{data}} a_l e^{i2\pi l\Delta kx}. \tag{96}$$

This model has several desirable properties. First, when no nontrivial *a priori* information is available, namely, $\mathcal{C}(\boldsymbol{\theta},x) = 1$, Eq. (96) automatically reduces to the conventional Fourier series model. This is very important, since the Fourier series model is indeed optimal in this case. On the other hand, if *a priori* information is available, this GS model provides a very effective means to incorporate this information into the final reconstruction. In the extreme case that $\mathcal{C}(\boldsymbol{\theta},x) = \rho(x)$, for example, the multiplicative Fourier series factor will be forced to unity by the data consistency constraint and $\rho_{GS}(x) = \rho(x)$ gives the exact reconstruction. In general, the new basis functions enable the GS model to converge faster than the Fourier series model (a_l decay faster than the original measured data) and, therefore, within a certain error bound, one can use far fewer terms to represent the spatial distributions of image information than are required by the Fourier series method, leading to a reduction of the truncation artifacts. In addition, improved spatial resolution is also possible since the GS reconstruction is determined by both the constraint function $\mathcal{C}(\boldsymbol{\theta},x)$ and the Fourier series factor $\Sigma_{l\in\mathcal{N}_{data}} a_l e^{i2\pi l\Delta kx}$. While the resolution of the latter is limited by the number of data points available, the resolution of $\mathcal{C}(\boldsymbol{\theta},x)$ is determined by the nature of *a priori* information. In many applications, $\mathcal{C}(\boldsymbol{\theta},x)$ comes from a high resolution reference image and superresolution reconstruction is obtained (144,247). For processing a given finite set of (symmetrically or asymmetrically sampled) Fourier data, $\mathcal{C}(\boldsymbol{\theta},x)$ has to be derived from the data themselves. For symmetrically sampled data,

it is suggested that $\mathcal{C}(\boldsymbol{\theta},x) = |\rho_{\mathrm{LPA}}(x)|$. In this way, $\boldsymbol{\theta}$ can be used to parametrize the edge locations as described in the previous section. Direct incorporation of sharp image edges into the basis functions in this way can help to remove Gibbs ringing from them. More importantly, when the LPA model is not appropriate or optimal, the GS model will automatically compensate by forcing the data consistency constraints. For achieving high computational efficiency, an approximation of such a constraint function can be obtained directly from the original Fourier image through the use of sigma-filtering (as discussed in the *Sigma Filter Extrapolation* section). The price to pay in this case is in the accuracy of the image edges, which are now limited by the Fourier pixel size. However, instead of directly transforming $\mathcal{C}(\boldsymbol{\theta},x)$ to get the missing high spatial frequency data, as is done in sigma-filter extrapolation methods, the GS model recovers the loss of image features in $\mathcal{C}(\boldsymbol{\theta},x)$ due to the filtering and, probably more importantly, it automatically reintroduces the missing phase information to $\mathcal{C}(\boldsymbol{\theta},x)$ through the multiplicative Fourier series factor to achieve the best possible data consistency between the extrapolated and original data. For asymmetrically sampled data, $\mathcal{C}(\boldsymbol{\theta},x)$ can be set to any of the partial-Fourier reconstructions obtained using the methods discussed in the *Phase-Constrained Reconstructions* section. These reconstructions give good image edge definitions, but may also contain various artifacts. It is suggested (138) that $\mathcal{C}(\boldsymbol{\theta},x)$ be simply set to the Hamming-filtered Fourier reconstruction with an appropriate edge enhancement. In this way, the GS model-based partial-Fourier reconstruction will be even faster than the iterative POCS methods.

Algorithm. Image reconstruction from finite-Fourier data using the GS model described by Eq. (96) consists of the following steps.

1. *Determination of the constraint function $\mathcal{C}(\boldsymbol{\theta},x)$:* In the symmetric data case, $\mathcal{C}(\boldsymbol{\theta},x)$ can be obtained using the LPA modeling algorithms described previously. An approximation of such a constraint function can be obtained more efficiently through the use of sigma-filtering (45,132). In this approach, a zero-padded Fourier reconstruction is first obtained with the desired digital resolution. The magnitude image is then sigma-filtered, with the same filter parameters as described in the *Sigma Filter Extrapolation* section, to remove spurious ringing and noise. Other edge-preserving filters, such as those discussed in (69,70,244,245) can also be used. This filtered image is then used as the constraint function $\mathcal{C}(\boldsymbol{\theta},x)$, providing the boundary constraints for the GS basis functions.

   In the asymmetric data case, the data $s(n\Delta k)$; $-n_0 \leq n < N$ are first zero-padded to $-N$, followed by Hamming-filtering on the first n_0 data points ($-n_0 \leq n < 0$) to roll off the sharp transition as described previously. The data is then Fourier transformed and the resultant magnitude image, $\rho_{\mathrm{DFT}}(x)$, is edge-enhanced through an exponential transform, $[\rho_{\mathrm{DFT}}(x)]^{\alpha}$ for $\alpha \sim 1.5$, to provide the necessary constraint function for the GS model.
2. *Solution of the series coefficients α_l:* After $\mathcal{C}(\boldsymbol{\theta},x)$ is available, the series coefficients are determined under the data consistency constraint

$$s(k) = \int_{-\infty}^{\infty} \left\{ \mathcal{C}(\boldsymbol{\theta},x) \sum_{l \in \mathcal{N}_{data}} a_l e^{i2\pi l \Delta k x} \right\} e^{-i2\pi k x} dx. \tag{97}$$

   Based on the convolution theorem of the Fourier transform, one has

$$s(n\Delta k) = \sum_{l \in \mathcal{N}_{data}} a_l s_0[(n - l)\Delta k], \tag{98}$$

where

$$s_0(n\Delta k) = \int_{-\infty}^{\infty} \mathcal{C}(\boldsymbol{\theta},x) e^{-i2\pi n\Delta k x} dx. \tag{99}$$

When the LPA model parameters are available for describing $\mathcal{C}(\boldsymbol{\theta},x)$, $s_0(n\Delta k)$ can be calculated analytically. Otherwise, $s_0(n\Delta k)$ can be obtained by applying a discrete Fourier transform to $\mathcal{C}(\boldsymbol{\theta},x)$. For example, if $\mathcal{C}(\boldsymbol{\theta},x)$ is the sigma-filtered magnitude image, the FFT will give $s_0(n\Delta k)$. Since $\mathcal{C}(\boldsymbol{\theta},x)$ is a real function, a dedicated FFT algorithm, such as the one in (225), can be used for improved computational efficiency.

In contrast to the usual linear prediction (autoregressive) formalism discussed earlier, Eq. (98) implies that the data $s(n\Delta k)$ is cross-linear predictable (cross-regressive) from $s_0(n\Delta k)$. This is a useful generalization of the classical LP theory, since the flexibility in selecting the basis signal $s_0(n\Delta k)$ will make $s(n\Delta k)$ exactly cross-linear predictable from $s_0(n\Delta k)$ to a finite order. Rewrite Eq. (98) in matrix form

$$\boldsymbol{H}\boldsymbol{a} = \boldsymbol{s} \tag{100}$$

then

$$\boldsymbol{H} = \begin{bmatrix} s_0(0) & s_0(-\Delta k) & \cdots & s_0[(-N_{data} + 1)\Delta k] \\ s_0(\Delta k) & s_0(0) & \cdots & s_0[(-N_{data} + 2)\Delta k] \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ s_0[(N_{data} - 1)\Delta k] & s_0[(N_{data} - 2)\Delta k] & \cdots & s_0(0) \end{bmatrix} \tag{101}$$

is a Toeplitz matrix (or block-Toeplitz matrix in the higher-dimensional cases). Therefore, the GS model coefficients $\boldsymbol{\alpha}$ can be found from Eq. (100) very efficiently using the Levinson algorithm (27,122). The matrix $\boldsymbol{H}$ is sometimes ill-conditioned. In this case, Eq. (100) needs to be regularized. A simple method is to perturb $\boldsymbol{H}$ with a diagonal matrix $\mu\boldsymbol{I}$. This modification on $\boldsymbol{H}$ corresponds to a new regularized GS model as

$$\tilde{\rho}_{GS}(x) = \mathcal{C}(\boldsymbol{\theta},x) \sum_{l \in \mathcal{N}_{data}} a_l e^{i2\pi l\Delta k x} + \mu \sum_{l \in \mathcal{N}_{data}} a_l e^{i2\pi l\Delta k x}, \tag{102}$$

in which the second part can absorb broad-band noise. Therefore, this regularized GS model is to be preferred for numerical stability and noise insensitivity. The regularization parameter μ can be chosen on the basis of the data signal-to-noise ratio and the amount of regularization required to stabilize $\boldsymbol{H}$, if that should be necessary. For example, μ can be chosen so that $s_0(0)/\mu \sim$ SNR.

3. *Generation of the GS image:* After $\mathcal{C}(\boldsymbol{\theta},x)$ and the a_l are available, evaluation of each pixel value directly from Eq. (96) will give the model reconstruction. Often a zero-filled FFT is applied to evaluate the Fourier series factor so that it has the same digital resolution as $\mathcal{C}(\boldsymbol{\theta},x)$. Sometimes, filtering on a_l may also be useful for removal of noise and residual ringing artifacts resulting from inaccuracy of the edge locations in $\mathcal{C}(\boldsymbol{\theta},x)$. The cross-LP equation (98) can be used to regenerate the missing high frequency data, followed by the conventional Fourier reconstruction to get the final image. This latter approach is often desirable, particularly for partial-Fourier reconstruction. This is because filtering on the GS coefficients and model regularization could lead to data inconsistency beyond the noise level, and data extrapolation using the model can reenforce the data consistency constraints in the final reconstruction.

Characteristics. The properties of the GS model for image reconstruction from finite-Fourier data are summarized as follows.

- *Flexibility:* The GS model is very flexible and effective for handling *a priori* information in a data consistent way. It is justified (145) that, if $\mathcal{C}(\boldsymbol{\theta},x)$ is well chosen, the GS model is asymptotically optimal in terms of the maximum cross-entropy principle, namely,

$$\rho_{GS}(x) \sim \max_{\{\rho(x)\}} \left\{ - \int_{\alpha}^{\alpha} \rho(x) \log \frac{\rho(x)}{\mathcal{C}(\boldsymbol{\theta},x)} dx \right\} . \tag{103}$$

  Note that the classical entropy constraint tends to force the final reconstruction to a global mean under the data limiters and is usually useful only for suppressing artifacts on the baseline (see the *Maximum Entropy Methods* section for more details). Under the cross-entropy constraint, the limiting global "mean" is $\mathcal{C}(\boldsymbol{\theta},x)$ rather than a constant. Therefore, noise and spurious features superimposed on $\mathcal{C}(\boldsymbol{\theta},x)$ can be removed in the final reconstruction, leaving only the true features evidenced in the measured data. This property of the cross-entropy principle has proved very useful for a variety of problems (209). By choosing the basis functions as the $|\rho_{LPA}(x)|$ modulated complex sinusoids, for example, this model bridges the gap between the LPA model (discussed in the previous section) and the classical Fourier series model. Since the LPA model is effective and efficient for characterizing sharp variations, while the Fourier series is optimal for representing smooth changes, the integration of these two features into a single, unified GS model is very desirable. Immediate consequences of this unification are faster convergence of the model and reduction of data truncation artifacts. Resolution improvement is also possible, depending on how $|\rho_{LPA}(x)|$ is determined. The accuracy of the image edge locations built into the basis functions will be retained in the final GS reconstruction. In some cases, it may be helpful to further generalize the GS model such that (145)

$$\rho_{GS}(x) = \mathcal{C}(\boldsymbol{\theta},x) \frac{\sum_{n=0}^{q} b_n e^{i2\pi n\Delta kx}}{1 + \sum_{n=1}^{p} a_n e^{i2\pi n\Delta kx}} = \mathcal{C}(\boldsymbol{\theta},x) \frac{B_q(e^{i2\pi\Delta kx})}{A_p(e^{i2\pi\Delta kx})} . \tag{104}$$

In this way, $\mathcal{C}(\boldsymbol{\theta},x)$ can be optimized for absorbing sharp edges, $A_p(e^{i2\pi\Delta kx})$ for spike features and $B_p(e^{i2\pi\Delta kx})$ for smooth changes.

The GS model can be applied to both symmetrically and asymmetrically sampled data. For asymmetric data reconstruction, this method offers a faster alternative to the Cuppen-POCS method discussed for partial-Fourier reconstruction.

- *Data consistency and stability:* The GS model reconstruction is, in principle, always data consistent, but model regularization and filtering of the series coefficients may lead to data inconsistency beyond the noise level. If that is the case, the data consistency constraint can be enforced by using the GS model for data extrapolation. This GS model-based extrapolation reconstruction is especially to be preferred over direct model reconstruction for asymmetric partial-Fourier data.

  The stability of the GS model is dependent on how the basis functions are chosen. For the basis functions described by Eq. (95), the regularized model in Eq. (102) is guaranteed to be stable if $\mathcal{C}(\boldsymbol{\theta},x)$ is a non-negative function.
- *Computational efficiency:* Fitting the GS model to the measured data for determining the model coefficients is very efficient. Especially when modeling is applied only to the phase-encoding or frequency-encoding direction, the computational time for finding the series coefficients is on the order of that required for two Fourier trans-

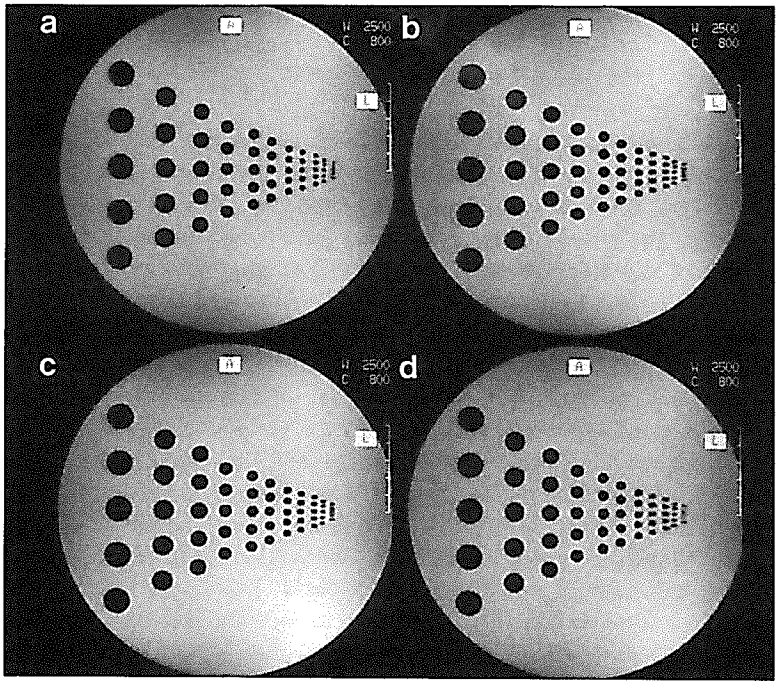

Figure 22. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a–b) the FFT with 256 and 128 symmetric data points, respectively; (c) the GS model with 128 symmetric data points and the constraint function (input image) from the sigma-filtered Fourier reconstruction; and (d) the same as (c) but the constraint function from the LPA model image. Note that residual ringing in (c) is reduced in (d) the edges are better defined in the basis functions through LPA modeling.

forms. However, computing the constraint function $\mathcal{C}(\boldsymbol{\theta},x)$ may cost some overhead time, depending on the application. In partial-Fourier reconstruction, the time required for calculating $\mathcal{C}(\boldsymbol{\theta},x)$ is that of one Fourier transform; for modeling truncated symmetric data, the time required would be that of one Fourier transform plus that for sigma filtering or that of LPA modeling fitting if superresolution edge definition is desired.

Figure 22 compares a set of reconstructions obtained using the GS model with those from the standard Fourier method. The images in Figs. 22a–b were reconstructed using the standard Fourier method with 256 and 128 data points, respectively. The GS images in Figs. 22c–d were reconstructed using 128 data points with the basis functions being constructed from the sigma-filtered Fourier image and the LPA model image, respectively. As can be seen, both GS images show significant suppression of the Gibbs ringing. Resolution improvement is also obvious for the GS image in Fig. 22d, as expected, since the edge constraints in the basis functions were obtained through LPA model fitting.

Figure 23 compares a set of partial-Fourier reconstructions obtained using the GS model with those obtained from the conventional Fourier method. As in Fig. 22, the images in Figs. 23a–b were reconstructed using the conventional Fourier method with 256 and 128 symmetric data points, respectively. The GS images in Figs. 23c–d were recon-

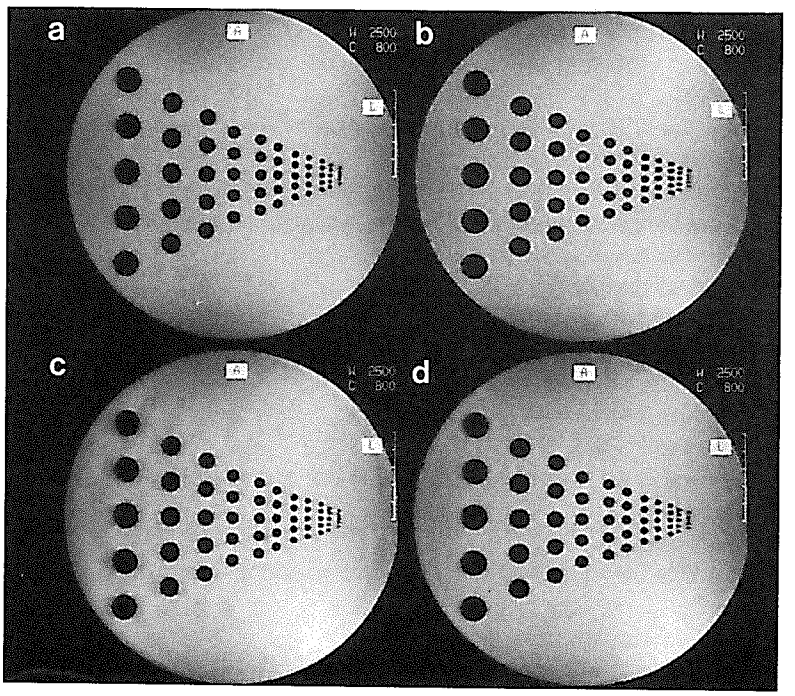

Figure 23. Reconstructions of a phantom image. The vertical direction is reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction is reconstructed using (a–b) the FFT with 256 and 128 symmetric data points, respectively; (c) the GS model-based symmetrization method with 128 asymmetric data points (n_0 = 16 points before the echo); and (d) the same as (c) but with n_0 = 32 points before the echo.

structed using 128 asymmetric data points with $n_0 = 16$ and 32 data points before the echo, respectively. The constraint function was obtained from the Hamming-filtered Fourier reconstruction as described in the *Theory* subsection. As can be seen, both GS partial-Fourier reconstructions give better resolution than the Fourier reconstruction from 128 symmetric data points. The GS images show much better suppression of ringing from asymmetric data truncation and also slightly better signal-to-noise ratio than the POCS images in Fig. 10.

Summary

In this section, we have discussed three parametric modeling methods for MR image reconstruction. These methods exploit some specific properties of the MR signals, such as linear predictability or piecewise continuity, to build an efficient parametric model for representing the image function so that the data truncation artifacts can be eliminated. A summary of the salient features of these methods is given below.

1. **The ARMA modeling method**
   - *Advantages:*
     - — reduced Gibbs ringing;
     - — improved resolution and SNR;
     - — moderately efficient;
     - — easy extension to 2-D (or 3-D) case, but with increased computational complexity.
     - — applicable to both symmetric and asymmetric data sets;
   - *Disadvantages:*
     - — selection of the model order (p,q), particularly p, is not trivial; the larger the p the better the resolution, but the worse the model stability;
     - — possible noise spikes;
     - — residual ringing from sharp edges.
2. **The LPA modeling method**
   - *Advantages:*
     - — reduction of Gibbs ringing;
     - — improved resolution for sharp edges;
     - — improved image SNR;
     - — applicable to both symmetric and asymmetric data sets;
   - *Disadvantages:*
     - — inefficient;
     - — possible loss of subtle image features when the direct model reconstruction is used;
     - — residual ringing from subtle edges;
     - — difficult extension to higher-dimensional cases.
3. **The GS modeling method**
   - *Advantages:*
     - — flexibility for using various constraints;
     - — computationally efficient;
     - — reduction of Gibbs ringing;
     - — possible improvement of image resolution and SNR;

— applicable to both symmetric and asymmetric data sets;
— easy extension to 2-D (or 3-D) case, but with increased computational complexity.

- *Disadvantages:*
  — determination of the basis functions may not be computationally efficient;
  — resolution improvement and ringing suppression depends on the basis functions;
  — residual ringing from inaccurate boundary constraints.

CONSTRAINED SPECTROSCOPIC IMAGING

Introduction

In the previous sections, we have focused on discussing anatomical imaging with data truncation in only one dimension. We now shift to reviewing a much more difficult problem — spectroscopic imaging — where data truncations take place normally in all spatial directions. Although constrained spectroscopic imaging is still in its infancy, a few methods have been proposed (98,136,141,142,229). We feel this review is timely since MR spectroscopic imaging (or spatially-resolved spectroscopy) is becoming increasingly important for metabolic imaging, but has encountered considerable difficulties in approaching the clinical realm because of the limited spatial resolution and SNR. Use of *a priori* information appears inevitable for overcoming this problem. We hope this preliminary discussion will stimulate more research in this direction.

The mathematical principle of MR spectroscopic imaging is simple. Like conventional anatomical imaging, we collect a set of data $s(\boldsymbol{k},t)$ in Fourier space and hope to invert the following Fourier integral equation to find the spectral function $\rho(\boldsymbol{x},f)$

$$s(\boldsymbol{k},t) = \int_{-\infty}^{\infty} \int_{\mathrm{FOV}} \rho(\boldsymbol{x},f) e^{-i2\pi(\boldsymbol{k}\boldsymbol{x}+ft)} d\boldsymbol{x} df, \tag{105}$$

where FOV is the field of view. As compared to the anatomical imaging equation in, Eq. (1), one may notice that a vector notation is used for the multi-dimensional frequency and space coordinate variables $\boldsymbol{k}$ and $\boldsymbol{x}$. In addition, a spectral frequency dimension is included in Eq. (105) for representing distributions of chemical shift information. While this new dimension of information is very important for biomedical research and clinical applications, it poses considerable technical difficulties if experiments are to provide enough data to obtain adequate spatial resolution (localization accuracy). This is because the freedom to use the time evolution course of the MR signals to encode spatial information has been lost, and all positional information has to be encoded through phase or frequency encoding by stepping the magnetic field gradients. Therefore, the number of phase-encoding (or frequency-encoding) steps available in practice is significantly limited, typically, $8 \times 8 \times 8$ in a 4-D (3 spatial + 1 spectral dimensions) experiment with an appropriate number of signal averagings. With the conventional Fourier reconstruction method [often called the Fourier CSI (chemical-shift imaging) method] (36), the resultant Gibbs ringing can cause serious spectral leakage and the pixel-limited resolution gives rise to significant partial-volume artifacts, both degrading the localization accuracy. One may naively suggest using extended phase-encodings with fewer or no signal averagings (for zero-mean

random noise, signal averaging is the most effective way for improving SNR), but this may lead to unacceptable degradation in SNR, which would render the improvement of localization meaningless (note that the data points collected near the origin of k-space contribute most to the SNR while peripheral data points contribute mainly to the spatial resolution). As we know from the *Fourier Reconstruction Method* section the loss of SNR efficiency in this way is irreversible with the Fourier method. Therefore, in order to achieve high spectral SNR, it is always beneficial to use a reduced set of phase-encodings with multiple signal averagings. One, then, has to use *a priori* information (constraints) for improving spatial localization. SLIM (Spectral Localization by IMaging) (98) is one such approach that utilizes structural information from high-resolution anatomical imaging to compensate for the lack of sufficient high spatial frequency phase-encoded spectroscopic data. There also exists a large body of literature on "spin-engineering" approaches that use volume-selective pulse sequences for improved spectral localization. These methods are usually useful for single voxel localization. For an excellent review, see (8). In this section, we will concentrate on only those postprocessing (reconstruction) methods making use of *a priori* constraints. These include the compartmental modeling method and the generalized series modeling method. Both methods are based on the SLIM principle of integration of anatomical imaging with spectroscopic imaging.

The SLIM Compartment Model

Theory. The SLIM technique is motivated by the fact that, in biological systems, water proton distribution is closely related to the distributions of other metabolites of interest, be they phosphorus metabolites or the proton/carbon-bearing metabolites. Particularly, the boundaries of the body, organs, tissue types, or lesions revealed by high-resolution water proton imaging often represent metabolic concentration discontinuities. By incorporating this *a priori* boundary information into the reconstruction of the spectral function $\rho(\boldsymbol{x},f)$, improved spatial resolution will result. The first approach to utilizing this boundary information is through the use of compartmental modeling. In this approach, the anatomical image is used to define a set of generalized voxels $\mathcal{D}_m$, called compartments, with arbitrary shapes and sizes to match exactly the biological regions of interest; and one wishes to reconstruct the average spectral distributions within individual compartments without inter-compartmental leakage, as shown in Fig. 24. This is in contrast to the conventional Fourier CSI method (36) where each voxel is a cube and inter-voxel leakage is inevitable because of the imperfect point spread function (the sinc function, see *Fourier Reconstruction Method* section for more details).

Mathematically, for an M-compartment model, the SLIM spectral function is given by

$$\rho_{\text{slim}}(\boldsymbol{x},f) = \sum_{m=1}^{M} \rho_m^{\text{slim}}(f), \tag{106}$$

where $\rho_m^{\text{slim}}(f)$ is an estimate of the desired average compartmental spectrum $\rho_m(f)$, namely

$$\rho_m^{\text{slim}}(f) \approx \rho_m(f) = \begin{cases} \dfrac{1}{V_m} \displaystyle\int_{\mathcal{D}_m} \rho(\boldsymbol{x},f)d\boldsymbol{x}, & \text{for } \boldsymbol{x} \in \mathcal{D}_m \\ 0, & \text{otherwise,} \end{cases} \tag{107}$$

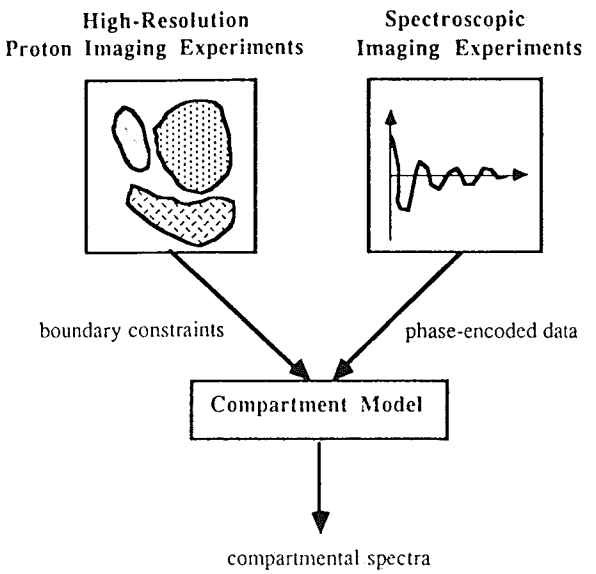

Figure 24. SLIM paradigm.

in which $\mathcal{D}_m$ is the compartmental geometry defined from the corresponding anatomical image and V_m is the volume of $\mathcal{D}_m$. Note that this definition of the SLIM spectral function is slightly different from the original one in (98), which is derived directly from a homogeneous compartment model. When this homogeneity assumption is valid, $\rho_m^{\text{slim}}(f)$ will be identical to the true spectral function $\rho_m(f)$ and can be found exactly. But, as shown below, the compartmental homogeneity is not a necessary condition for SLIM to derive the average compartmental spectral distribution with any desired accuracy. One may also notice that this SLIM compartment model is a multi-dimensional extension of the zeroth-order LPA model already discussed, but does not require the parametrization of locations and thus eliminates the need to solve a nonlinear optimization problem for finding the edges from the spectroscopic data.

Algorithm. With the SLIM model in Eqs. (106) and (107), the spectral reconstruction problem reduces to finding the compartmental spectrum $\rho_m^{\text{slim}}(f)$. Hu et al. (98) suggested fitting the SLIM model directly to the data such that

$$\min_{\{\rho_m^{\text{slim}}(f)\}} \sum_{n=1}^{N_{data}} \left| s(\boldsymbol{k}_n,t) - \int_{\text{FOV}} \left[\sum_{m=1}^{M} \int_{-\infty}^{\infty} \rho_m^{\text{slim}}(f) e^{-i2\pi ft} dt \right] e^{-i2\pi \boldsymbol{k}_n \boldsymbol{x}} d\boldsymbol{x} \right|^2. \tag{108}$$

In this way, the boundary information can be forced into the reconstruction for improved localization. Under the model constraint in Eq. (106), one obtains

$$\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1M} \\ g_{21} & g_{22} & \cdots & g_{2M} \\ \cdot & & & \cdot \\ \cdot & \cdots & \cdots & \cdot \\ \cdot & & & \cdot \\ g_{N_{data}1} & g_{N_{data}2} & \cdots & g_{N_{data}M} \end{bmatrix} \begin{bmatrix} c_1^{\text{slim}}(t) \\ c_2^{\text{slim}}(t) \\ \cdot \\ \cdot \\ \cdot \\ c_M^{\text{slim}}(t) \end{bmatrix} \approx \begin{bmatrix} s(\boldsymbol{k}_1,t) \\ s(\boldsymbol{k}_2,t) \\ \cdot \\ \cdot \\ \cdot \\ s(\boldsymbol{k}_{N_{data}},t) \end{bmatrix}, \tag{109}$$

where g_{nm} is the phase dispersal factor of the nth phase-encoding wave vector across the m^{th} compartment

$$g_{nm} = \int_{\mathcal{D}_m} e^{-i2\pi k_n x} dx, \tag{110}$$

and $c_m^{\text{slim}}(t)$ is the compartmental spectroscopic signal related to $\rho_m^{\text{slim}}(f)$ by

$$c_m^{\text{slim}}(t) = \int_{-\infty}^{\infty} \rho_m^{\text{slim}}(f) e^{-i2\pi ft} df. \tag{111}$$

In vector notation, Eq. (109) can be written simply as

$$\boldsymbol{G}\boldsymbol{c}^{\text{slim}}(t) \approx \boldsymbol{s}(t) \tag{112}$$

where $\boldsymbol{G}$ is frequently called the SLIM geometry matrix for the obvious reason, and the least squares solution for $\boldsymbol{c}^{\text{slim}}(t)$ is given by

$$\boldsymbol{c}^{\text{slim}}(t) = (\boldsymbol{G}^H\boldsymbol{G})^{-1}\boldsymbol{G}^H\boldsymbol{s}(t) = \boldsymbol{G}^+\boldsymbol{s}(t). \tag{113}$$

Interestingly, as a comparison, if the boundary constraints were applied after normal Fourier reconstruction, one would get (140)

$$\boldsymbol{c}^{\text{ft}}(t) = \begin{bmatrix} \frac{1}{V_1} & & & \\ & \frac{1}{V_2} & & \\ & & \cdots & \\ & & & \frac{1}{V_m} \end{bmatrix} \boldsymbol{G}^H\boldsymbol{s}(t) \tag{114}$$

The weakness of this approach is that any artifacts in the Fourier reconstruction will pass down to the reconstructed compartmental spectra.

In summary, reconstruction of the SLIM spectral function $\rho_{\text{slim}}(\boldsymbol{x},f)$ includes the following major steps:

1. Segmentation of the anatomical images into a set of generalized voxels $\{\mathcal{D}_m; m = 1, \cdots, M\}$. This should be done based on the desired regions of interest. It may require a complicated image segmentation procedure if full automation is desired.
2. Calculation of the SLIM geometry matrix $\boldsymbol{G}$. The phase dispersal factor g_{nm} defined in Eq. (110) can be computed through numerical integration.
3. Solution of the SLIM equation (109). Usually, the number of phase-encoding steps N_{data} is chosen to be larger than the number of compartments M. The least squares solution in Eq. (113) can be obtained using a variety of numerical methods (73). Since M is usually a small number, (10 or less, for example), a SVD-based procedure can be used to gain numerical stability, and to get the singularity condition measure of $\boldsymbol{G}$, $\kappa(\boldsymbol{G})$, as a side-product without computationally over-burdening the procedure.

4. Inverse Fourier transform $c_m^{\text{slim}}(t)$ to get $\rho_m^{\text{slim}}(f)$. Often, correction of baseline and phase artifacts is needed before the spectra can be used for spectral quantitation (143).

Characteristics. In this subsection, we will examine the sensitivity of $\rho_{\text{slim}}(x,f)$ to noise and compartmental inhomogeneities (modeling offsets) and discuss the possible advantages of the SLIM method.

Effects of measurement noise — In practical applications, random measurement noise is always present. Assume that the measured data $\hat{s}(t)$ are perturbed by $\boldsymbol{n}(t)$, and that as a consequence $\boldsymbol{c}^{\text{slim}}(t)$ is in error by the vector $\Delta\boldsymbol{c}^{\text{slim}}(t)$. Then we have

$$\boldsymbol{G}[\boldsymbol{c}^{\text{slim}}(t) + \Delta\boldsymbol{c}^{\text{slim}}(t)] = \boldsymbol{s}(t) + \boldsymbol{n}(t),$$

and the SLIM compartmental spectroscopic signal becomes

$$\hat{\boldsymbol{c}}^{\text{slim}}(t) = \boldsymbol{c}^{\text{slim}}(t) + \boldsymbol{G}^{+}\boldsymbol{n}(t). \tag{115}$$

The following conclusions can be drawn directly from Eq. (115).

- Since $\Delta\boldsymbol{c}^{\text{slim}}(t) = \boldsymbol{G}^{+}\boldsymbol{n}(t)$ is independent of the underlying spectral structures $\boldsymbol{c}(t)$, random noise, no matter what the data SNR is, will not cause inter-compartmental spectral leakage. This is understandable since random noise, in principle, has no spatial origin and is independent of the phase-encoding gradients so that its presence does not violate the compartmental homogeneity assumption (139).
- Different compartments have different noise sensitivity depending on their size and shape and the phase-encoding vectors. Specifically,

$$\Delta c_m^{\text{slim}}(t) = \sum_{l=1}^{N_{data}} g_{ml}^{+} n_l(t), \tag{116}$$

and the variance of $\Delta c_m^{\text{slim}}(t)$ is

$$\text{Var}\{\Delta c_m^{\text{slim}}(t)\} = \sum_{l=1}^{N_{data}} |g_{ml}^{+}|^2 \sigma_{data}^2 = \|\boldsymbol{g}_m^{+}\|^2 \sigma_{data}^2, \tag{117}$$

where $\sigma_{data} = \sqrt{\text{Var}\{n_l(t)\}}$ and $\boldsymbol{g}_m^{+} = (g_{m1}^{+}, \cdots, g_{mN_{data}}^{+})$ is the m^{th} row vector of the pseudo-inverse matrix $\boldsymbol{G}^{+}$. It can be justified that $\|\boldsymbol{g}_m^{+}\| \propto 1/V_m$ (140) and, thus, the larger a compartment, the lower is its noise sensitivity. One can also find that the overall noise sensitivity is bounded by (73)

$$\frac{\|\Delta\boldsymbol{c}^{\text{slim}}(t)\|}{\|\boldsymbol{c}^{\text{slim}}(t)\|} \leq \kappa(\boldsymbol{G}) \frac{\|\boldsymbol{n}(t)\|}{\|\boldsymbol{s}(t)\|}, \tag{118}$$

where $\kappa(\boldsymbol{G})$ is the condition number of $\boldsymbol{G}$, which under the spectral norm is given by $\kappa(\boldsymbol{G}) = \lambda_{\max}/\lambda_{\min}$, with $\lambda_{\max}$ and $\lambda_{\min}$ being the maximum and minimum singular

values of $\boldsymbol{G}$, respectively. In general, $\kappa(\boldsymbol{G}) \geq \sqrt{V_{max}/V_{min}}$ with V_{max} and V_{min} being the largest and smallest volume of all the SLIM compartments, which suggests that partitioning the field of view into compartments of comparable sizes would help to reduce the overall *worst-case* noise sensitivity. In addition, since k-space sampling in SLIM is not restricted by the Nyquist criterion, one has the freedom to choose the phase-encoding vectors $\{\boldsymbol{k}_n\}$ to achieve optimal noise sensitivity for a particular compartment (of interest) by minimizing $\|\boldsymbol{g}_m^+\|$ with respect to $\boldsymbol{k}_n$, at the expense of the noise sensitivity of other non-interesting compartments.

- For a fixed set of phase-encoded measurements, the Fourier inversion method gives the lowest noise sensitivity (but not necessarily the best SNR, since localization of the signal is poor). From Eq. (114), one has

$$\mathrm{Var}\{\Delta c_m^{\mathrm{ft}}(t)\} = \frac{1}{V_m^2} \sum_{n=1}^{N_{data}} |g_{nm}^*|^2 \sigma_{data}^2 = \|\boldsymbol{g}_m^*\|^2 \sigma_{data}^2 / V_m^2. \tag{119}$$

The noise amplification factor for SLIM can then be defined as

$$\mathcal{A}_m = \sqrt{\frac{\mathrm{Var}\{\Delta c_m^{\mathrm{slim}}(t)\}}{\mathrm{Var}\{\Delta c_m^{\mathrm{ft}}(t)\}}} = \frac{V_m \|\boldsymbol{g}_m^+\|}{\|\boldsymbol{g}_m^*\|}, \tag{120}$$

which quickly approaches 1 as $\boldsymbol{G}$ becomes stable (140), as shown in Fig. 26.

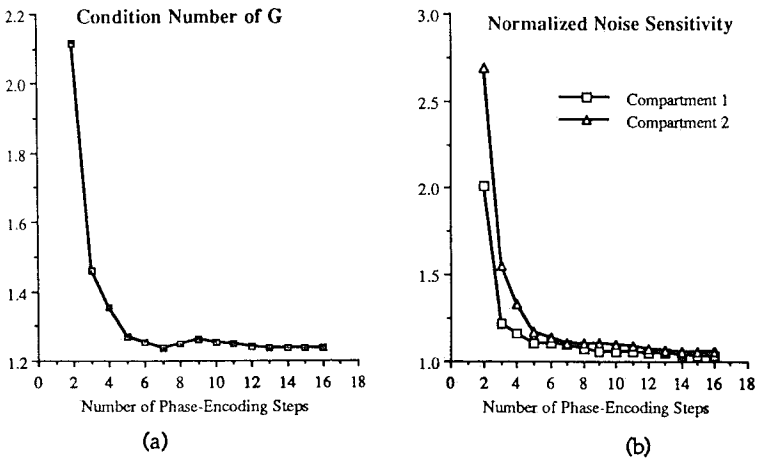

Figure 25. SLIM noise sensitivity simulation results based on a 1d phantom (compartment 1 defined as $-w_1 < x < 0$ and compartment 2 as $0 < x < w_2$, where w_1 = 3/8 of the FOV and w_2 = 1/4 of the FOV): (a) the condition number of the SLIM geometry matrix G; note that it takes 4 phase-encoding steps to stabilize the SLIM equation because of the small size of the second compartment; and (b) the SLIM noise sensitivity normalized against the Fourier method where compartmental constraints are applied after the reconstruction. As can be seen, SLIM will not amplify noise when an appropriate number of phase-encoding steps is used so that G is well conditioned.

Effects of inhomogeneity errors — Compartmental inhomogeneities arise in the SLIM model when any compartments are inaccurately defined. This is especially the case when several non-interesting compartments are lumped together, and/or when there exist magnetic field inhomogeneities or metabolic inhomogeneities not revealed by the water proton images. The effects of compartmental inhomogeneities on SLIM are thoroughly examined in (140). The main results are summarized as follows.

- It is proved that

$$\lim_{N_{data}\to\infty} \rho_m^{\text{slim}}(f) = \rho_m(f), \tag{121}$$

which indicates that the SLIM solutions are asymptotically *unbiased* no matter what compartmental inhomogeneities exist. Namely, for any application, the SLIM results can be made as precise as one wishes by sufficiently over-determining the SLIM equation. This is one of the most interesting and important theoretical results about SLIM. It assures that, irrespective of the enforced boundary constraints, the SLIM compartment model is not restricted to modeling compartmentally-homogeneous spectral functions only, but, coupled with the least-squares principle, it can handle compartmental inhomogeneities of any kinds without systematic biases.

- For a finite N_{data}, SLIM will suffer a truncation error in the form of inter-compartmental spectral leakage. Interestingly, comparing Eq. (113) with Eq. (114), this error comes from the same source as the Fourier method — truncation of each Fourier series element in $\boldsymbol{G}^H\boldsymbol{s}(t)$,[16] but is compensated for in SLIM by the off-diagonal elements of the $(\boldsymbol{G}^H\boldsymbol{G})^{-1}$ matrix. This compensation is complete when all the compartments are homogeneous, leading to a leakage-free reconstruction. In the presence of compartmental inhomogeneities, this correction is incomplete, resulting in residual intercompartmental leakage in the final SLIM reconstruction. Generalizing the usual point spread function (PSF) concept for measuring leakage from one point source to another, the SLIM leakage can be characterized through a contribution function (CF), $\mathcal{P}_m^{\text{slim}}(x)$, which measures the contribution from a point source at x to a compartment m. Ideally, $\mathcal{P}_m^{\text{slim}}(x)$ should be a constant for $\boldsymbol{x} \in \mathcal{D}_m$ and zero elsewhere, meaning that the localized spectrum for compartment m receives contributions from compartment m only. For SLIM (140,242),

$$\mathcal{P}_m^{\text{slim}}(\boldsymbol{x}) = \sum_n g_{mn}^{+} e^{-i2\pi \boldsymbol{k}_n \boldsymbol{x}}, \tag{122}$$

in contrast to that of the Fourier methods given by

$$\mathcal{P}_m^{\text{ft}}(\boldsymbol{x}) = \frac{1}{V_m} \sum_n g_{nm}^{*} e^{-i2\pi \boldsymbol{k}_n \boldsymbol{x}}. \tag{123}$$

[16]The m^{th} element of $\boldsymbol{G}^H\boldsymbol{s}(t)$ is $\int_{\mathcal{D}_m} \sum_{n=1}^{N_{data}} s(\boldsymbol{k}_n, t) e^{i2\pi \boldsymbol{k}_n \boldsymbol{x}} d\boldsymbol{x}$.

Spectral leakage from compartment n to compartment m is given by (140)

$$\mathcal{L}_{mn} = \int_{\mathcal{D}_n} \rho(\boldsymbol{x},f)\, \mathcal{P}_m(\boldsymbol{x})d\boldsymbol{x}. \tag{124}$$

This formula is useful for both SLIM and the Fourier method if one replaces $\mathcal{P}_m$ with the corresponding CF. It is found (140) that $\mathcal{P}_m^{\text{slim}}(x)$ has several advantages over $\mathcal{P}_m^{\text{ft}}(x)$ for reducing the inter-compartmental spectral leakage. (a) No spectral leakage will come from any homogeneous compartments, i.e., $\mathcal{L}_{mn} = 0$ if $\rho(\boldsymbol{x},f)$ is homogeneous for $\boldsymbol{x} \in \mathcal{D}_n$. (b) For any inhomogeneous compartment, it is found by decomposing the inhomogeneities into different harmonics that (140): (a) the zero-order (constant) component of $\rho(\boldsymbol{x},f)$, $\boldsymbol{x} \in \mathcal{D}_n$ will not cause any spectral leakage, even if there exist sharp inter-compartmental transitions; this often amounts to a significant reduction of the Gibbs spectral leakage associated with the Fourier method; (b) leakage from higher-order components with spatial frequencies well above the highest phase-encoding wave number $\|\boldsymbol{k}\|_{\max}$ will be negligible (see Fig. 26), which implies that heterogeneities in a biological system found at "microscopic" levels will not be a problem to SLIM; and (c) low-frequency components, such as field inhomogeneities in B_0 or B_1, will cause inter-compartmental spectral leakage if they are unmodeled. However, since truncation error from these smooth

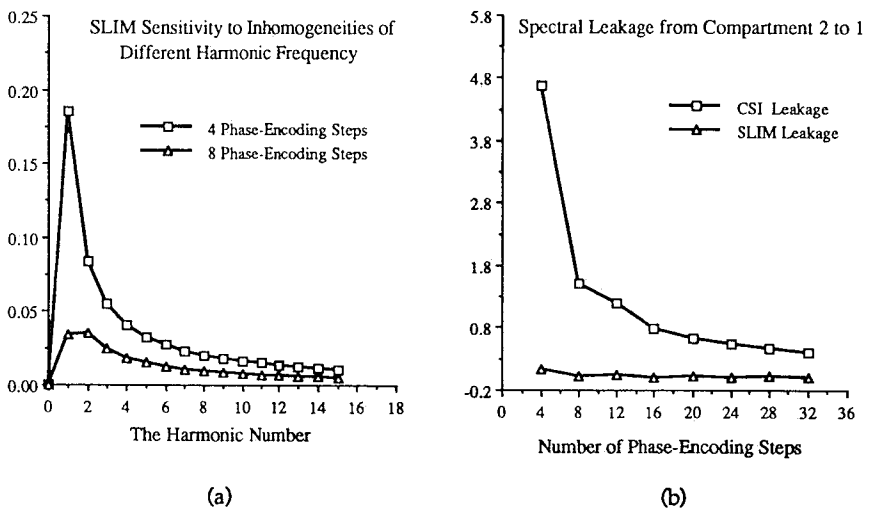

Figure 26. Simulation results of SLIM sensitivity to compartmental inhomogeneities based on a 1-d phantom (same as the one in Fig. 25 but with $w_1 = w_2 = 1/4$ of the FOV). Compartment 1 is homogeneous with concentration $\rho_1(x) = 1$ and compartment 2 has a concentration inhomogeneity. The plots in (a) show the spectral leakage to compartment 1 when the concentration in compartment 2, $\rho_2(x)$, is set to $1 + \sin(2\pi n f_0 x)$ where $f_0 = 1/w_2$ is the basic harmonic frequency. As can be seen, SLIM is sensitive to inhomogeneities over a limited range of frequencies. The plots in (b) show the total spectral leakage of SLIM and CSI to compartment 1 when an inhomogeneity consisting of 32 harmonics with a unit amplitude is added to compartment 2. Note that SLIM has a much faster convergence than CSI and the leakage with SLIM becomes negligible when $\boldsymbol{G}$ becomes stable (see the condition number plot in Fig. 25).

components is small, they may not cause significant spectral leakage if not amplified in the SLIM solution process, [which is guaranteed as long as the SLIM equation is not ill-conditioned (139)].

Imaging efficiency — It is clear from Eq. (109) that, if the phase-encoding vectors are nontrivially chosen, the SLIM spectral function $\rho_{slim}(\boldsymbol{x},f)$ can be obtained exactly, provided the number of phase-encoding measurements, N_{data}, is not less than the number of compartments, M. This method, therefore, can provide exact spatial localization (super-resolution) with the minimal number of phase-encoding steps when the homogeneous compartment model holds. This time efficiency is one of the most important features of SLIM. When compartmental inhomogeneities exist, the optimal efficiency of SLIM is lost, and one needs to use more phase-encoding steps than the number of compartments defined to secure a reliable solution. For a given set of phase-encoded measurements, as long as the compartments are not ill-defined (i.e., the superimposed boundaries represent some kinds of spectral discontinuities) and the SLIM equation is not ill-conditioned (or singular), SLIM will yield a better measurement of the average compartmental spectral intensity than the conventional Fourier method. On the other hand, given a SLIM model, the minimal number of phase-encoded measurements and the optimal phase-encoding wave vectors $\{\boldsymbol{k}_l\}$ can be found by minimizing $\kappa(\boldsymbol{G})$ with respect to $\{\boldsymbol{k}_l\}$. As long as the resultant $\kappa(\boldsymbol{G})$ is on the order of $\sqrt{V_{max}/V_{min}}$ the SLIM result will be reliable. It is found (140) that the number of phase-encoding steps required to stabilize $\boldsymbol{G}$ is much less than that needed by the Fourier method for achieving a comparable localization accuracy (spatial resolution).

The Generalized SLIM Model

Theory. As discussed in the previous section, the SLIM compartment model is useful for obtaining average compartmental spectra. However, the phase-encoded data may also contain information about the intra-compartmental variations of the spectral distribution. This information is lost as a least squares fitting error in the SLIM calculation process. The generalized SLIM (GSLIM) model is developed to salvage this useful information so as to reveal intra-compartmental spectral distribution under the same boundary constraints (136). Mathematically, the GSLIM model is based on the generalized series modeling concept, detailed in the *Parametric Modeling Methods* section. It expresses the spectral function in terms of a generalized series as

$$\rho(\boldsymbol{x},f) = \sum_n a_n(f)\varphi_n(\boldsymbol{x},f), \tag{125}$$

where the basis functions are now built as

$$\varphi_n(\boldsymbol{x},f) = \rho_{slim}(\boldsymbol{x},f)e^{i2\pi k_n x}, \tag{126}$$

and the GSLIM spectral function becomes

$$\rho_{gslim}(\boldsymbol{x},f) = \rho_{slim}(\boldsymbol{x},f)\sum_n a_n(f)e^{i2\pi k_n x}. \tag{127}$$

Besides the general useful properties discussed previously, this model offers a unique way to unify compartmental spectral localization and high resolution multi-voxel Fourier imaging. This unification is important since it allows both *a priori* information and experimental data to be utilized in a most effective and data-consistent way. In principle, $\rho_{\text{gslim}}(\boldsymbol{x},f)$ can provide a localized spectrum from any spatial location, $\boldsymbol{x}$, as high resolution multi-voxel imaging does. The boundary information built into the GSLIM basis functions [through $\rho_{\text{slim}}(\boldsymbol{x},f)$] can help to register the spectral information more accurately, thus improving spatial resolution and reducing spectral leakage. For compartmental localization, on the other hand, the GSLIM model can absorb compartmental inhomogeneities through the multiplicative Fourier series factor, thereby extending the ability of the simple compartment model to handle hidden inhomogeneities. Specifically, the compartmental spectra from the GSLIM model are now given by

$$\rho_m^{\text{gslim}}(f) = \frac{1}{V_m}\int_{\mathcal{D}_m} \rho_{\text{gslim}}(\boldsymbol{x},f)d\boldsymbol{x}\ \rho_m^{\text{slim}}(f)\sum_n g_{mn}^* a_n(f)/\text{V}_m \tag{128}$$

in which the second series factor can correct for some SLIM spectral leakage from higher-order compartmental inhomogeneities present in $\rho_m^{\text{slim}}(f)$ due to data truncation.

Algorithm. Calculation of the GSLIM spectral function $\rho_{\text{gslim}}(\boldsymbol{x},f)$ requires finding the SLIM spectral function $\rho_{\text{slim}}(\boldsymbol{x},f)$ and the series coefficients $a_n(f)$. $\rho_{\text{slim}}(\boldsymbol{x},f)$ can be obtained following the procedure outlined in the previous section, and the series coefficients $a_m(f)$ should then be determined under the data consistency constraint such that

$$s(\boldsymbol{k},t) = \int_{-\infty}^{\infty}\int_{\text{FOV}} \{\rho_{\text{slim}}(\boldsymbol{x},f)\sum_n a_n(f)e^{i2\pi\boldsymbol{k}_n\boldsymbol{x}}\}e^{-i2\pi(\boldsymbol{k}\boldsymbol{x}+ft)}d\boldsymbol{x}df. \tag{129}$$

Applying the inverse Fourier transform with respect to t to both sides gives

$$\tilde{s}(\boldsymbol{k},f) = \int_{\text{FOV}} \left\{\rho_{\text{slim}}(\boldsymbol{x},f)\sum_n a_n(f)e^{i2\pi\boldsymbol{k}_n\boldsymbol{x}}\right\}e^{-i2\pi\boldsymbol{k}\boldsymbol{x}}d\boldsymbol{x}. \tag{130}$$

In practice, $\tilde{s}(\boldsymbol{k},f)$ can be obtained using the discrete Fourier transform, since enough data points are collected along the time axis. Based on the convolution theorem of the Fourier transform, the above equation can be simplified to

$$\tilde{s}(\boldsymbol{k},f) = \sum_n a_n(f)H(\boldsymbol{k}-\boldsymbol{k}_n,f), \tag{131}$$

where the convolution kernel is

$$H(\boldsymbol{k},f) = \int_{\mathcal{D}} \rho_{\text{slim}}(\boldsymbol{x},f)e^{-i2\pi\boldsymbol{k}\boldsymbol{x}}d\boldsymbol{x}. \tag{132}$$

In vector notation, Eq. (131) becomes

$$\boldsymbol{H}(f)\boldsymbol{a}(f) = \tilde{\boldsymbol{s}}(f). \tag{133}$$

This equation often needs to be regularized as

$$[\boldsymbol{H}(f) + \mu\boldsymbol{I}]\boldsymbol{a}(f) = \tilde{\boldsymbol{s}}(f), \tag{134}$$

to gain numerical stability as described in the *Generalized Series Model* sections. The $\boldsymbol{a}(f)$ can be found from this regularized equation using a general linear system solver. Since this system has a Toeplitz structure in the one-dimensional case, and a block-Toeplitz structure in higher-dimensional cases, specialized algorithms, such as those in (27,113,122) can be employed to achieve higher computational efficiency. In summary, GSLIM reconstruction includes the following major steps:

1. Extract boundary constraints from the high resolution anatomical images using an appropriate image segmentation procedure (as in SLIM).
2. Calculate the SLIM spectral function $\rho_{slim}(\boldsymbol{x},f)$, as described in the previous section. Phase correction is often needed on $\rho_{slim}(\boldsymbol{x},f)$ before it is built into the GSLIM basis functions.
3. Apply the temporal Fourier transformation to individual phase-encoded FIDs, $s(\boldsymbol{k},t)$, to get $\tilde{s}(\boldsymbol{k},f)$. Often, removal of phase and baseline artifacts, resulting from phase-encoding delays and the presence of broad spectral components, is necessary in this step (141,142).
4. Calculate the GS coefficients $\boldsymbol{a}(f)$ from Eq. (134).
5. Calculate the GSLIM average compartmental spectra from Eq. (128). To reconstruct a metabolic map for a particular spectral line at chemical shift frequency f_0 with spectral linewidth f_w, a multi-dimensional FFT is applied to the GS coefficients $\boldsymbol{a}(f)$ at each frequency over the frequency range $|f - f_0| < f_w/2$ to reveal the intra-compartmental variations. Zero-padding is often used so that the FFT reconstruction has the same digital resolution as the boundary constraints (or the anatomical images); mild filtering may also be useful to roll off the sharp transitions between the padded zeros and the original data. To improve SNR, the spectral function $\rho_{slim}(\boldsymbol{x},f)$ thus obtained from Eq. (127) is integrated over the frequency range $|f - f_0| < f_w/2$ to give the desired metabolic map.

Characteristics. The GSLIM model has several desirable properties for making effective use of anatomical information for spectroscopic imaging. In the trivial case that $\rho_{slim}(\boldsymbol{x},f) = 1$, i.e., when no anatomical constraints are superimposed, it automatically reduces to the conventional Fourier series representation. This property of GSLIM is important, since the Fourier series model is optimal when no *a priori* information is available. In the other extreme, if $\rho_{slim}(\boldsymbol{x},f)$ can precisely model $\rho(\boldsymbol{x},f)$, the Fourier series factor will be forced to be 1 by the data consistency requirement, and GSLIM becomes identical to SLIM. In practice, some significant *a priori* information can always be extracted from the high resolution proton image and be built into $\rho_{slim}(\boldsymbol{x},f)$. Even if the SLIM compartment model is inadequate, (due to compartmental inhomogeneities, for example), the Fourier series factor will provide a mechanism to reveal and accommodate such intra-compartmental variations (in concentration, composition, or magnetic field), and the SLIM inhomogeneity leakage can, therefore, be reduced and improved localization ob-

tained. The main properties of the GSLIM method for constrained spectroscopic imaging are summarized as follows:

- It unifies SLIM compartmental localization and high resolution multi-voxel imaging.
- *A priori* boundary information can be effectively incorporated into the reconstruction of the desired spectral functions without violating data consistency.
- Outstanding spectral discontinuities represented by the *a priori* boundary constraints can help to eliminate data truncation leakage from them; hidden compartmental spectral variations evidenced in the spectroscopic data will be revealed without being subject to artificial compartmental homogeneity assumptions.
- It is applicable to any dimensionality, but is computationally rather expensive for 4-D data sets.

Figure 27 shows a set of reconstructions from a two-dimensional model (spatial axis along the horizontal direction and spectral frequency axis along the vertical direction) with three compartments of different spectral properties. The Fourier reconstruction, Fig. 27c, obtained with 12 phase-encoding steps displays serious spectral leakage along the spatial axis (also see the profile along the frequency axis). To test the capability of GSLIM to handle compartmental inhomogeneities, two compartments on the right hand side were lumped together as a single inhomogeneous compartment. Under this "wrong" structural constraint, the SLIM reconstruction, Fig. 27c, gives the average spectrum from the inhomogeneous compartment, as expected. Spectral leakage resulting from the inhomogeneous compartment can also be seen, but is much less serious than the Gibbs spectral leakage with the CSI reconstruction. With the GSLIM model under the same structural constraint, the spectral leakage is significantly reduced, thanks to its ability to accommodate compartmental inhomogeneities. Additionally, the spatial distributions of spectral information within the compartments are reasonably well reconstructed, as desired.

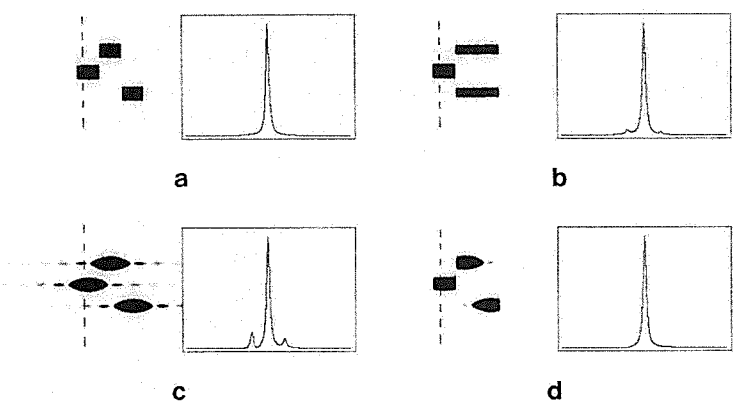

Figure 27. (a) the model (spatial axis along the horizontal direction and the frequency axis along the vertical direction) with three compartments of different spectral properties; (b) SLIM reconstruction with the rightmost two compartments being lumped as an inhomogeneous compartment; (c) Fourier CSI reconstruction with 12 phase-encoding steps; and (d) GSLIM reconstruction under the same structural constraint, as in (b).

Summary

In this section, we have discussed two preliminary but rather promising methods for constrained spectroscopic imaging. These methods are based on the novel idea of using high resolution anatomical information to constrain spectroscopic distributions originated by Lauterbur and Levin for overcoming the inefficiency problem with the normal CSI method. The SLIM compartmental modeling method developed by Hu et al. permits spectral localization from generalized voxels with natural shapes and sizes to match biological regions of interest from a reduced set of phase-encoding measurements. The GSLIM method developed by Liang et al. further extends SLIM with an additional step for compartmental imaging. A summary of the salient features of these methods is given below.

1. **The SLIM compartmental modeling method**
   - *Advantages:*
     - — Improved imaging efficiency (and thus, SNR efficiency);
     - — Minimized partial-volume artifacts;
     - — Reduced Gibbs spectral leakage;
     - — Applicable to volume-selective data acquired with surface coils or selective pulses;
     - — Permissible to use nonuniform sampling k-space for further improved compartmental spectral localization;
     - — Computationally more efficient than CSI processing;
     - — Useful for compartmental localization of spectral information from ROIs with arbitrary shapes and sizes.
   - *Disadvantages:*
     - — Possible residual data truncation artifacts (inter-compartmental spectral leakage) from compartmental inhomogeneities;
     - — Possible loss of information available for revealing compartmental inhomogeneities.
2. **The GSLIM modeling method**
   - *Advantages:*
     - — Unification of SLIM compartmental localization and high resolution CSI multi-voxel imaging.
     - — *A priori* boundary information can be effectively incorporated into the reconstruction of the desired spectral function without violating data consistency and without using the artificial compartmental homogeneity assumption.
     - — Minimized partial-volume artifacts;
     - — Capable of reconstructing compartmental spectral variations (inhomogeneities);
     - — Useful for both compartmental spectral localization and high resolution metabolic imaging.
   - *Disadvantages:*
     - — Computationally more expensive than the CSI method;
     - — Requirement of data preprocessing for phase and baseline corrections before GSLIM reconstruction.

PRACTICAL EXAMPLES, DISCUSSIONS, AND CONCLUSION

Introduction

Up to this point, we have focused on the theoretical discussion of each reconstruction algorithm with its main features being demonstrated by reconstructions from a single phantom data set. In this section, we will provide the reader with a more detailed description of the behavior of these reconstruction techniques based on three clinical data sets.

The first example is a sagittal head slice image. The raw data was collected using a gradient-echo sequence with $T_E = 10$ms and $T_R = 200$ms. This data set is a challenge to most of these reconstruction methods because it has rapid phase changes (Fig. A2 in the Appendix) and is a good example for demonstrating the effects of rapid phase variations on partial-Fourier reconstruction. The second example is a T_2-weighted transverse head slice image. The raw data was collected using a spin-echo sequence with $T_E = 90$ms and $T_R = 2500$ms. The last example is a T_1-weighted spin-echo sagittal lumbar spine image collected using $T_E = 15$ms and $T_R = 500$ms. The last two sets should be excellent testing grounds for the reconstruction methods; if they cannot fare well in these examples the methods will be of little practical use.

For all of the anatomical imaging examples, the initial raw data matrices contain 256 $\times$ 256 complex data points. Each data set was Fourier transformed along the phase encoding (vertical) direction using all of the available 256 points. The resultant semi-raw data set was then truncated along the read-out (horizontal) direction to the desired sizes before a given reconstruction algorithm was applied to get the final image. In the spectroscopic imaging case, a 4-D ^{31}P phase-encoded data set acquired from a human head was used for comparison of SLIM and GSLIM reconstruction with CSI reconstruction.

Throughout this chapter, the reader will find comments describing the nature and scope of image artifacts associated with a given reconstruction. Since there is no single best approach to quantitate the quality of these images, we decided to compare them based on the type of observations that a radiologist would have made under the same conditions. Thus, all of the images presented here were commented on by at least one radiologist. While great effort has been made to make the descriptions as precise and representative as possible, the comparison of various methods here is by no means the most rigorous, and deserves further research. Therefore, no rating or ranking of an individual method is given. It is the authors' opinion that selecting a particular method for a specific application often involves consideration of many parameters and is best judged by the informed user.

Partial-Fourier Reconstructions

Partial-Fourier imaging is now becoming increasingly popular, as it allows very short echo times to be used for minimizing T_2^* signal loss and spin dephasing due to flow effects. Our aim in this section is to examine the practical utility of various reconstruction methods reviewed in this paper. For each example, we vary the degree of data asymmetry, but fix the total number of data points so that the total imaging time is the same if it is applied in the phase-encoding direction. For simplicity, we denote the total number of data points as $(n_0 + N)$ with n_0 being the number of data points before the echo. Also, a standard Fourier image reconstructed from full symmetric echos is included for comparison.

Margosian reconstructions. The Margosian sagittal head reconstruction with (16 + 112) points is shown in Fig. 28b. This image contains severe artifacts, especially near the pituitary gland, the mouth cavity and the spinal cord, where large phase variations exist (Fig. A2). Intensity voids are evident at different locations, which can easily be confused with regions of legitimate low image intensity. When 32 points before the echo are used, the Margosian reconstruction, Fig. 29b, shows some improvement, but the artifacts remain significant. This example confirms the earlier theoretical prediction that a good phase estimate does not guarantee a good Margosian's reconstruction if phase variations are significant. It is, therefore expected that the Margosian method will have only limited utility in gradient-echo imaging.

In spin-echo imaging applications, the Margosian method produces satisfactory results because the phase variations are almost negligible (see Figs. A3 and A4). For example, for the transverse head data set, the Margosian reconstructions display no apparent artifacts with both 16 and 32 points before the echo. The image resolution is very good and the sulci and the narrow structures in the white matter are defined reasonably well. The main drawback of these images is poor signal-to-noise ratio as compared with the Fourier reconstruction from full echos (Figs. 30a and 31a). The same observations can be made about the Margosian reconstructions from the lumbar spine data set shown in Figs. 32b and 33b.

Cuppen-POCS Reconstructions. Figure 28c shows the POCS image of the sagittal head slice reconstructed using (16 + 112) points. As can be seen, the image artifacts have been significantly reduced compared to the Margosian counterpart (Fig. 28b). However, some artifacts are still present in the form of intensity loss and contour deformations due to the poor phase constraints. When the number of data points before the echo is increased to 32, the phase estimate improves accordingly. The POCS reconstruction under these phase constraints, Fig. 29c, shows no intensity voids in the mouth cavity, and the spatial extent of the small artifacts near the pituitary gland and in the vicinity of the spinal cord has also been reduced. Although some small intensity voids are still present at the air/skin interface where high frequency phase variations exist and are not picked up by the phase constraints, this example demonstrates that the POCS method can handle large phase variations reasonably well in gradient-echo imaging applications. For spin-echo data sets, the POCS reconstructions, Figs. 30c–33c, are similar to the corresponding Margosian reconstructions with a minor improvement in image signal-to-noise ratio.

GS reconstructions. The GS image of the sagittal head reconstructed using (16 + 112) points is shown in Fig. 28d. This image shows a clear improvement over the corresponding Margosian and POCS images in contrast and in the severity of image artifacts. For example, a considerably smoother behavior can be seen near the pituitary gland and along the skin/air interface. Also, the image contrast between the white matter and the rest of the tissue is somewhat closer to that of the Fourier image from full echos than it is in the other two images. This suggests that the GS method is better able to handle large phase variations than the other two methods, especially when the number of data points before the echo (n_0) is small. When n_0 is increased to 32, the GS reconstruction, Fig. 29d, improves accordingly, but with only minor improvements over the corresponding POCS reconstruction in Fig. 29c. For the spin-echo data sets, the GS reconstructions, Figs. 30d–33d, are very close to the corresponding POCS reconstruction, except in Fig. 31d,

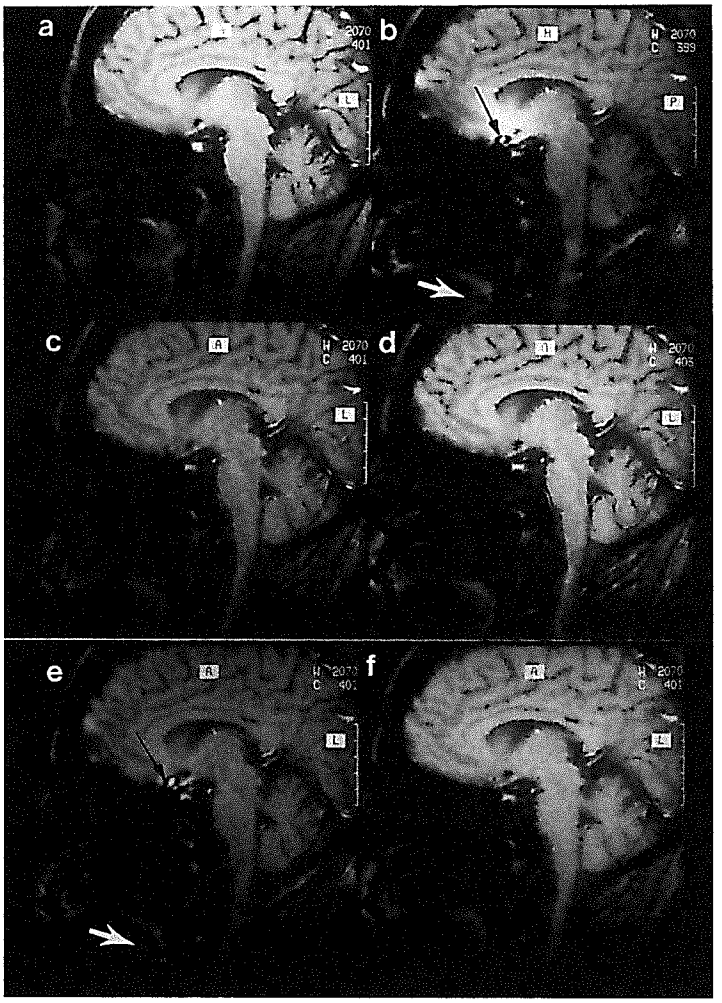

Figure 28. Reconstructions of a sagittal head slice. The vertical direction was reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction was reconstructed (a) from 256 symmetric data points using the FFT; and from 128 asymmetric data points, $-16 \leq n < 112$, using: (b) the Margosian method; (c) the Cuppen-POCS method; (d) the GS method; (e) the TERA method; and (f) the LPA method. The POCS image was reconstructed with 4 iterations; the GS image was reconstructed using a Hamming-filtered Fourier reconstruction as the initial input for constructing the basis functions; the TERA image was reconstructed using the Margosian reconstruction as the initial input for creating the pseudo data set for ARMA model fitting. Note the phase artifacts in the Margosian and TERA images, and the improvement of the GS image over the POCS image.

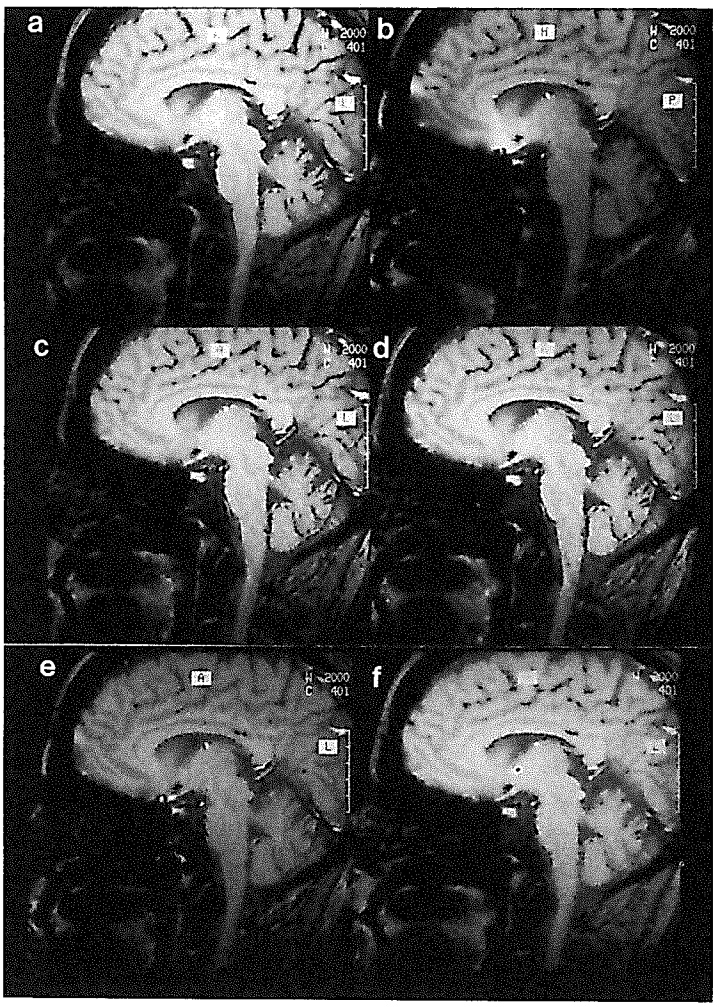

Figure 29. The same as Fig. 28, except that 32 points before the echo, $-32 \leq n < 96$, were used to reconstruct the partial-Fourier images in (b)–(f). All images are improved in image quality compared to those in Fig. 28, since the phase behavior is better defined by the central low frequency data. Note that the phase artifacts in the Margosian (b) and TERA (e) reconstructions remain significant; the improvement of the GS image over the POCS image is less obvious.

where sigma-filtering on the GS coefficients leads to a better signal-to-noise ratio but with minor resolution loss.

TERA reconstructions. The ARMA model-based reconstructions using the TERA algorithm described previously are theoretically expected to be a direct improvement over the Margosian reconstructions. However, for the sagittal gradient-echo head data sets (Figs. 28 and 29), the TERA images, Figs. 28e and 29e, suffer from the same phase artifacts as the Margosian images. Therefore, unlike the POCS and GS methods, the TERA algorithm

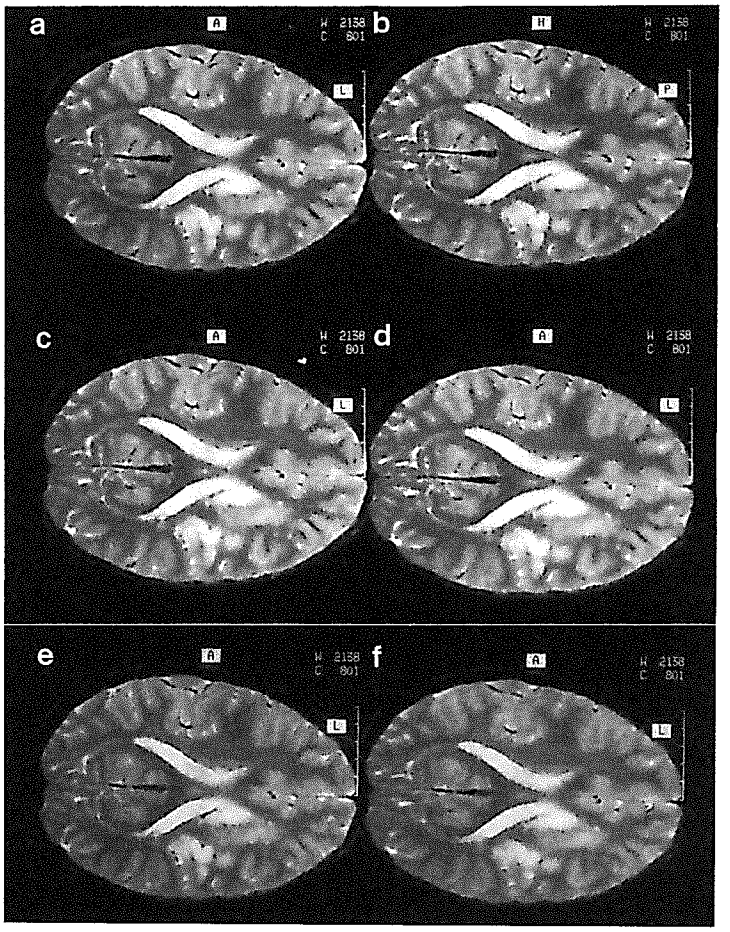

Figure 30. Reconstructions of a transverse head slice. All the images were reconstructed as described in Fig. 28. Because of the lack of significant phase variations in this spin-echo data set, the phase-constrained methods [Margosian (b) and POCS (c)] apparently perform very well, except a slight loss in signal-to-noise with the Margosian image. The modeling methods [GS (d), TERA (e), and LPA (f)] show no noticeable improvement in image quality, except some resolution loss with the LPA image.

as it stands does not extend the capability of Margosian's method for handling the large phase variations commonly encountered in gradient-echo imaging applications. For the spin-echo data sets, the TERA images show a slight resolution improvement over all other partial-Fourier images presented. This resolution improvement is particularly evident at the end of the spinal cord as well as at the CSF boundaries with the spinal cord and the sixth vertebra in Fig. 32e.

LPA reconstructions. The LPA image of the sagittal head slice reconstructed using (16 + 112) points is shown in Fig. 32f. This image has a contrast resembling that of the Fourier reconstruction from full echos of Fig. 28a. Some phase artifacts are still present near the

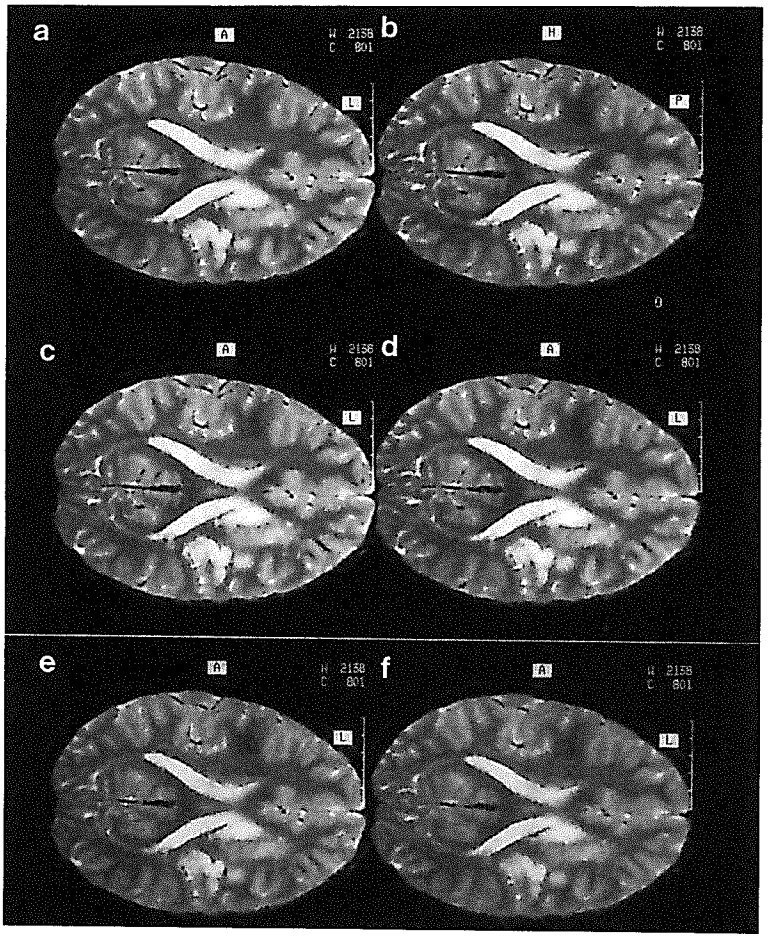

Figure 31. The same as Fig. 30, except that 32 points before the echo, $-32 \leq n < 96$, were used to reconstruct the partial-Fourier images in (b)–(f), but no significant improvement in image quality is observed.

pituitary gland and along the skin air interface but are less serious than in all other partial-Fourier images. These artifacts are reduced when the number of data points before the echo is increased to 32, but at the expense of a slightly higher level of resolution loss, as is clear in the cerebellum and in the posterior region of the neck. For the spin-echo data sets, the LPA reconstructions, Figs. 30f–33f, show a similar loss of resolution. Since LPA modeling does not reintroduce noise during the extrapolation step, it gives the best signal-to-noise ratio among all of the partial-Fourier images.

Summary

Margosian method: Computationally efficient; a good alternative to the conventional or filtered Fourier methods when there are no significant phase variations across the image, as in spin-echo imaging applications; new reconstruction artifacts in regions with *large* phase variations; signal-to-noise lower than any other methods being considered.

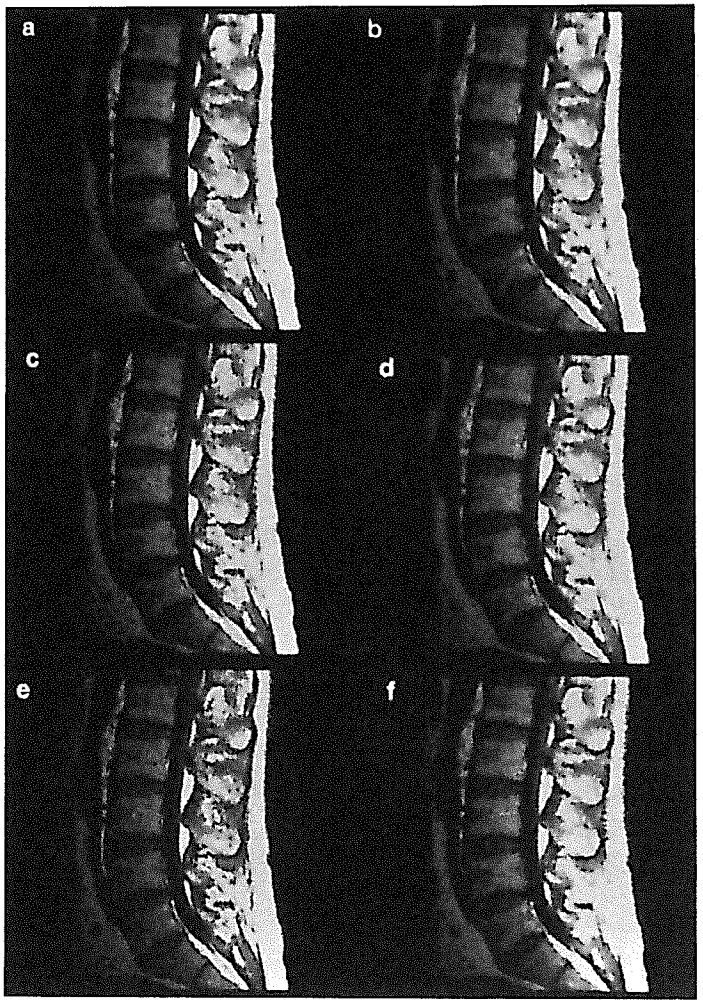

Figure 32. Reconstructions of a sagittal lumbar spine slice. All the images were reconstructed as described in Fig. 28. The relative quality of the partial-Fourier images is similar to that in Fig. 30, because both data sets have similar phase behavior.

Cuppen-POCS method: Computationally less efficient than Margosian's method; can handle large phase variations and yields exact reconstruction when the phase estimation is precise; possible reconstruction artifacts (ringing and signal cancellation) in regions with *high frequency* phase variations; lower signal-to-noise relative to the full Fourier transform reconstruction, but slightly better than the Margosian reconstructions.

GS method: Computationally less efficient than Margosian's method but faster than the Cuppen-POCS method; slightly better performance than Cuppen-POCS method, especially when very short echo times are used (large degree of data asymmetry); signal-to-noise comparable to the Cuppen-POCS method.

TERA method: Computationally less efficient than the above methods; slightly better performance than the Margosian method, but still unable to handle large phase variations;

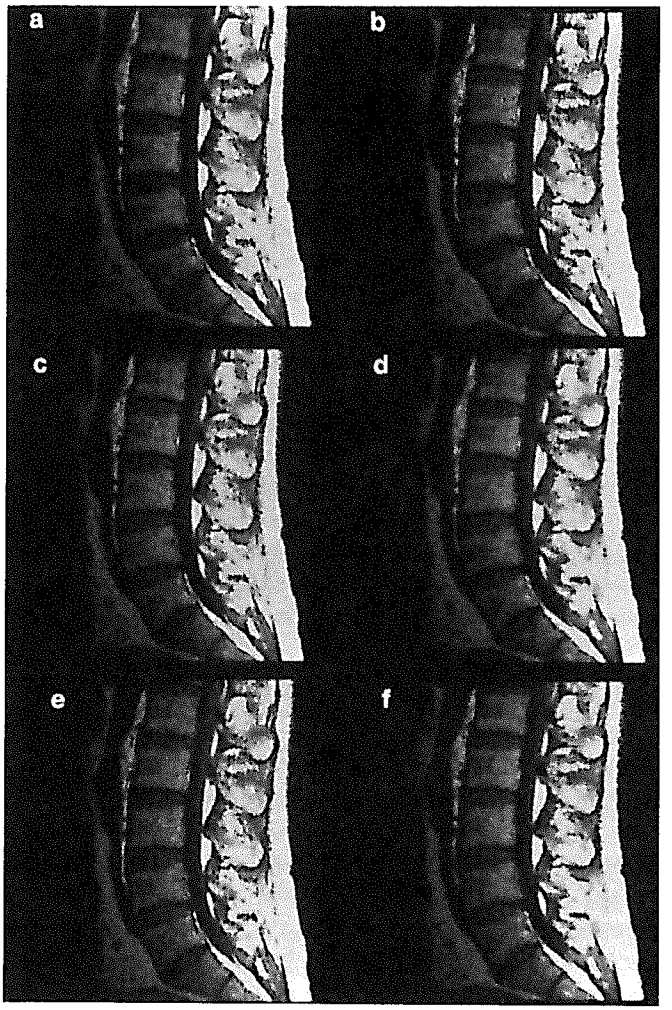

Figure 33. The same as Fig. 32 except that 32 points before the echo, $-32 \leq n < 96$, were used to reconstruct the partial-Fourier images in (b)–(f). Slight resolution loss can be observed as compared with the images in Fig. 32 because less high spatial frequency information is available in this data set. Note also the blurring in image (d) due to a sigma-filtering on the GS coefficients.

slightly better resolution than both the Cuppen-POCS method and the GS method; signal-to-noise comparable to Margosian's method.

LPA method: Computationally least efficient; good handling of phase errors; lowest resolution but the best signal-to-noise.

Truncated-Fourier Reconstructions

In truncated-Fourier imaging, the measured echo signal is symmetric and there are no phase-related artifacts such as those occur in partial-Fourier imaging. However, because

of the lack of high spatial-frequency data, the ability of a reconstruction method to reduce Gibbs' ringing and to improve resolution is most important. This issue is addressed below for each method, based on reconstructions from 96 and 128 data points. In order for the reader to better appreciate the quality of the new reconstructions, conventional Fourier reconstructions from both the full and the truncated data sets are also presented for comparison.

Sigma filter reconstruction. The SF image of the sagittal head slice reconstructed using 96 data points is presented in Fig. 34c. As compared with the direct Fourier reconstruction in Fig. 34b, this image shows superb ringing suppression but with a noticeable resolution loss. When the number of data points is increased to 128, a much improved SF reconstruction is obtained, Fig. 35c. The Gibbs ringing is now completely removed and the resolution loss is negligible. This somewhat dramatic improvement in the performance of the algorithm when compared with the 96 point reconstruction of Fig. 34c is due to the higher resolution present in the 128 point Fourier transform image that serves as its initial input. This higher resolution image allows the extrapolation step of the algorithm to produce more reliable high spatial-frequency data since the sharp discontinuities in the spin density profile are now more easily distinguished. Similar observations can be made about the SF reconstructions from the transverse head and lumbar spine data sets presented in Figs. 36c–39c.

GS reconstructions. The GS reconstructions are presented in Figs. 34d–39d. The basis functions were constructed based on the sigma-filtered result of a zero-padded Fourier reconstruction. No super-resolution is, therefore, expected for the GS reconstructions. As compared with the corresponding SF reconstructions, however, the GS reconstructions show somewhat better resolution and contrast because of better data consistency between the extrapolated data and the measured data. For example, in Figs. 34 and 35, the anatomical structures in the posterior arch of the C_1 vertebral body and at the interface between the spine and the esophagus are much better defined in the GS reconstructions than in the SF reconstructions.

TERA reconstructions. The TERA reconstructions are presented in Figs. 34e–39e. These images show some of the superresolution capabilities of the ARMA model. For example, in Figs. 34e and 35e, structures inside the cerebellum and in the vicinity of the pituitary gland are much better defined than in the corresponding zero-padded Fourier reconstructions; for the transverse head, some regions inside the white matter are defined with a similar clarity to the 256 point Fourier reconstruction. However, some artifacts can also be observed in some of the TERA reconstructions. For example, in Fig. 35e, there is some leftover ringing in the spinal cord as well as edge distortions inside the mouth cavity and the skin air interface of the forehead.

LPA reconstructions. The LPA reconstructions are presented in Figs. 34f–39f. These images show very good suppression of the Gibbs ringing. Resolution improvement, however, is feature-dependent. For some large, sharp image features, such as blood vessels, the LPA reconstructions show the sharpest edges, but for some subtle and narrow

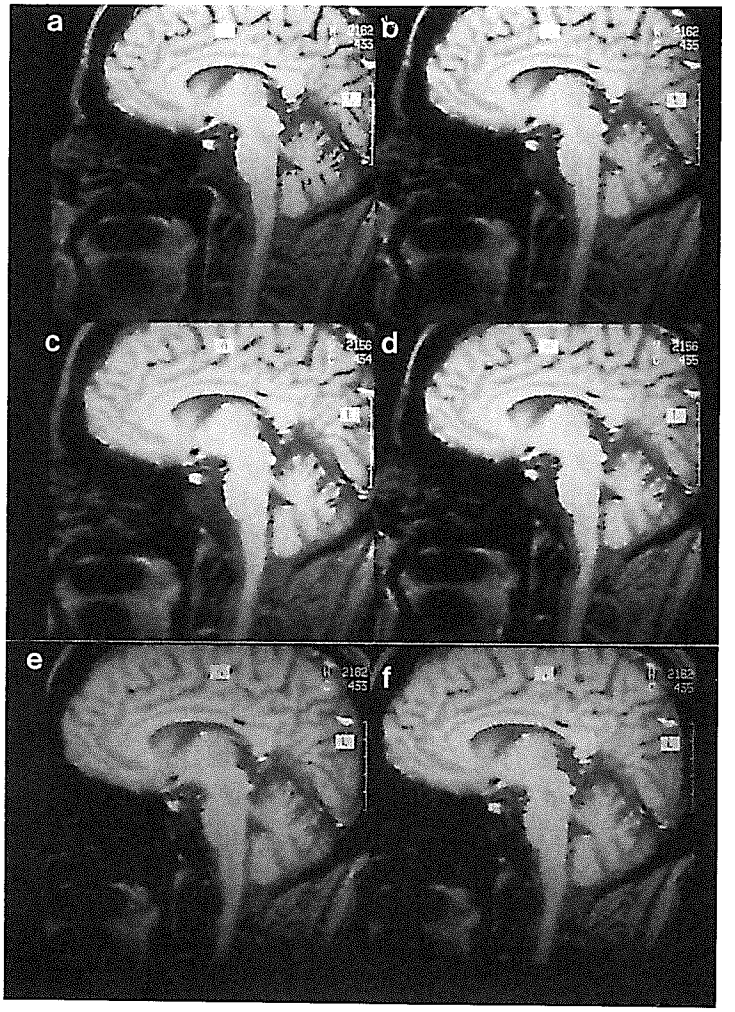

Figure 34. Reconstructions of a sagittal head slice. The vertical direction was reconstructed using the Fourier method from full echos of 256 data points. The horizontal direction was reconstructed (a) from 256 symmetric data points using the FFT; and from 96 asymmetric data points using: (b) the FFT; (c) the SF method; (d) the GS method; (e) the TERA method; and (f) the LPA method. The SF image was obtained using Amartur-Haacke's modified SF approach with 3 iterations; the GS image was reconstructed with the basis functions being built from (b) after sigma filtering; the TERA image was reconstructed with an AR order $p = 16$; and the LPA image was reconstructed using a SVD-based LP algorithm. All these images show good ringing suppression with the TERA image giving the best resolution.

image features, edge distortions are also evident, as can be seen in the posterior arch of the C_1 vertebral body and over the C_2 and C_3 vertebral bodies in the sagittal head slice images in Figs. 34 and 35. The boxy appearance of the gray and white matter in the LPA reconstructions in Figs. 36f and 37f are also due to the loss of resolution. This behavior of

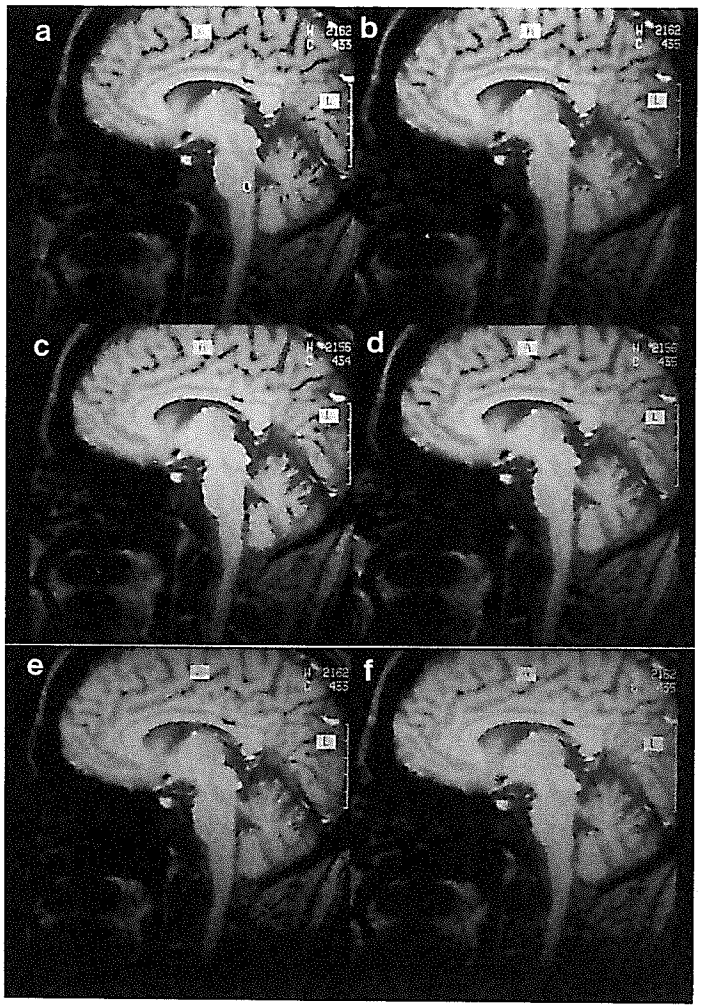

Figure 35. The same as Fig. 34 except that 128 data points were used for reconstruction of the images in (b)–(f). For the TERA image, the AR model order is $p = 20$. These images have better resolution than those in Fig. 34 because more high frequency data were available, but the syrinx-like artifacts are still present in the spinal cord in the Fourier image (b) and the TERA image (f).

the LPA reconstructions is largely because the existing LPA reconstruction algorithms fail to find the optimal estimate of the edge locations; especially when applied to high-pass filtered data, the enhanced noise variance leads to biases in the estimated edges of the subtle image features.

Summary

SF method: Computationally efficient; very good suppression of Gibbs ringing; slight resolution loss but no new reconstruction artifacts; signal-to-noise similar to the Fourier

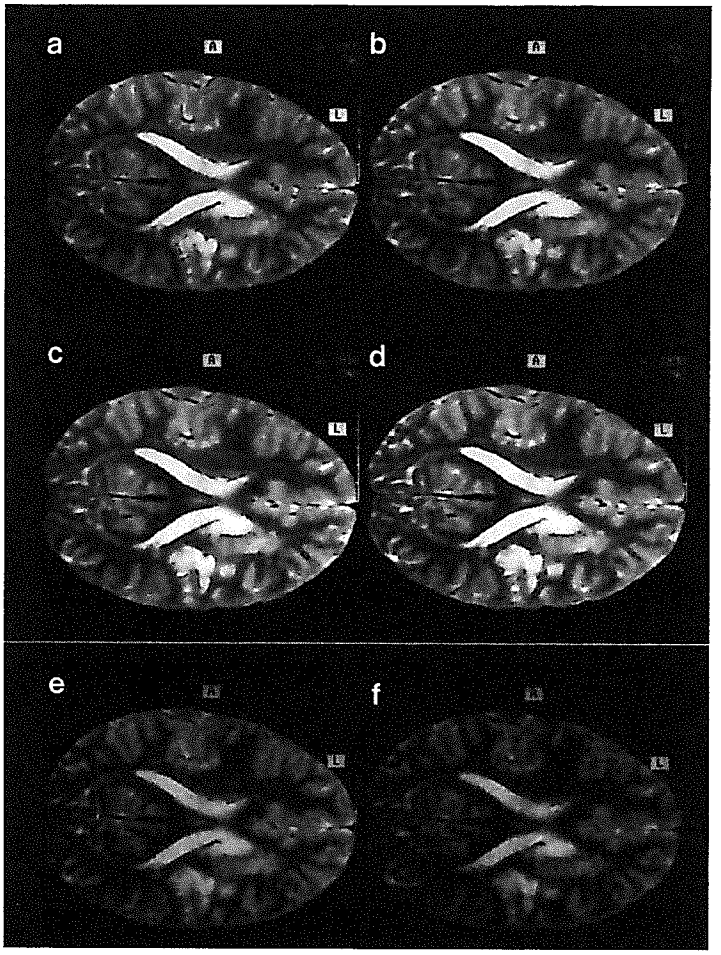

Figure 36. Reconstructions of a transverse head slice. All the images were reconstructed as described in Fig. 34. Because of the lack of sharp transitions in this image, the Gibbs ringing in the Fourier image (b) is not obvious, and other reconstructions display only minor improvement.

image from truncated data; a good alternative to the conventional Fourier method, especially for low levels of data truncation.

GS method: Computational efficiency comparable to the sigma-filter method (faster than Amartur-Haacke's iterative method); slightly better resolution; similar signal-to-noise; no new reconstruction artifacts.

TERA method: Computationally more expensive than the sigma-filter method; best resolution of all; signal-to-noise comparable to that of the Fourier image from the truncated data; some residual Gibbs' ringing and possible noise spikes.

LPA method: Computationally least efficient; very good suppression of Gibbs' ringing; improved resolution for large, high intensity image features; signal-to-noise comparable to that of the Fourier image from the truncated data; resolution improvements are not

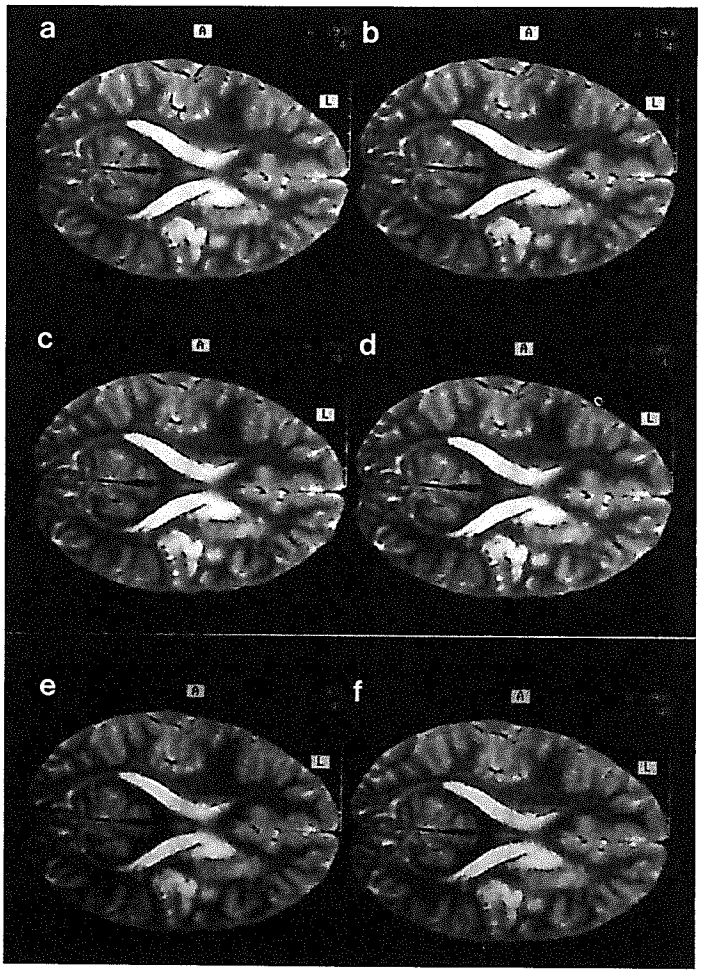

Figure 37. The same as Fig. 36, except that 128 data points were used for reconstruction of the images in (b)–(f), leading to slightly better resolution.

always obtained in practice, as they are dependent on the object sizes and the signal-to-noise in the input data.

Spectroscopic Reconstruction

Figure 40 shows the results of a SLIM analysis on a ^{31}P data set of $4 \times 4 \times 4$ phase encodings obtained from the head of a normal volunteer at 1.5T on a Siemens Magnetom using a 17 cm Helmholtz coil. As can be seen, SLIM gives good localization of the ^{31}P signals. The different PCr concentration in the cerebellum (with brain stem) and cerebrum revealed by this result is consistent with an early study in human cadavers (88).

A GSLIM reconstruction of the PCr metabolic map from these data is shown in Fig. 41b. The GSLIM maps show excellent localization as compared to the corresponding

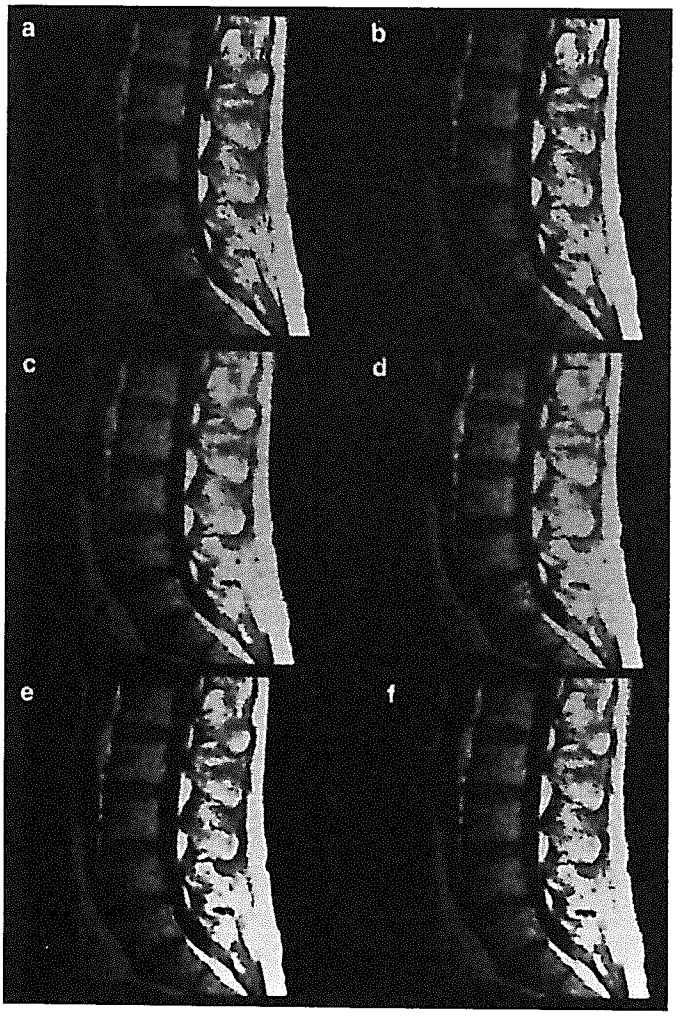

Figure 38. Reconstructions of a sagittal lumbar spine slice. All the images were reconstructed as described in Fig. 34. Unlike in Fig. 36 the truncated Fourier image shows significant Gibbs ringing running down the spinal cord. The LPA reconstruction gives the best ringing suppression and image sharpness.

Fourier reconstruction in Fig. 41a. In addition to reconstructing the biological metabolic distribution, the GSLIM method reveals nicely the compartmental variations due to the modulation of the inhomogeneous B_1 field and coil sensitivity, as is shown in Fig. 42. Further rigorous investigation of these techniques is necessary to evaluate their quantitative accuracy, which is beyond the scope of this review.

Conclusions

The Fourier transform has long been the mainstay of image reconstruction methods in MRI because of its speed and robustness. Nevertheless, it fails to give the optimal SNR

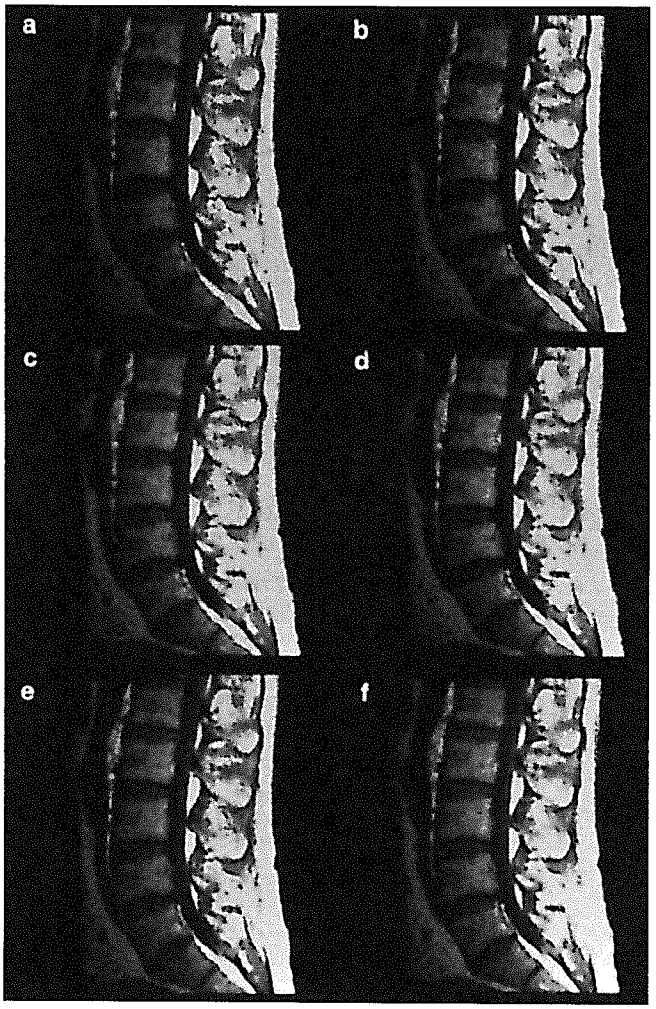

Figure 39. The same as Fig. 38, except that 128 data points were used for reconstruction of the images in (b)–(f), leading to reduced ringing and slightly better resolution.

and resolution for practical imaging data, so alternate reconstruction methods are being sought to give better performance. We have discussed a variety of these methods specific to MR (and sometimes with applications outside the MR domain) in an attempt to make the reader aware of the potential improvements possible. Although information cannot be created by smart mathematical manipulations, *a priori* constraints, if used properly, will result in a better image reconstruction technique than the conventional Fourier method. The improvement in performance is often manifested by higher resolution and reduced Gibbs ringing, as shown by various images presented here. These new techniques will, therefore, make it possible to shorten imaging times without sacrificing image quality. They also allow more flexible data acquisition (pulse sequence design) for MR angiography, for example, through the use of asymmetric echos. One of the practical limitations of constrained reconstruction methods is longer processing times, but this can be over-

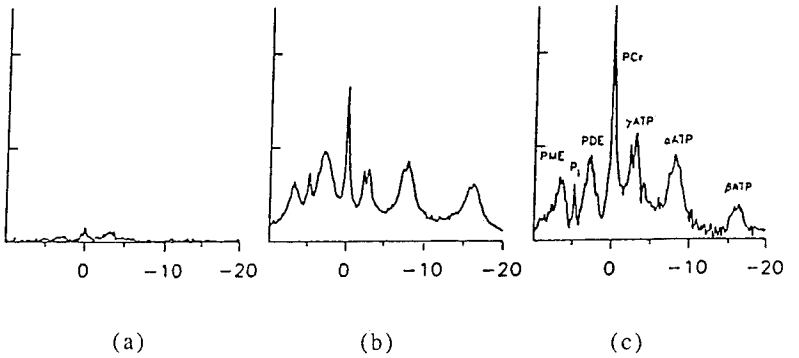

Figure 40. SLIM localized ^{31}P spectra from a data set of $4 \times 4 \times 4$ phase encodings obtained from the head of a normal volunteer at 1.5T on a Siemens Magnetom using a 17 cm Helmholtz coil. The vertical axis is adjusted so that peak integrals represent relative concentrations (not yet corrected for T_1 saturation) within the 3 volumes: (a) non-brain volume, i.e., skull, muscle, scalp, nasal cavity, etc., within the coil; (b) predominantly cerebrum; and (c) cerebellum with brain stem.

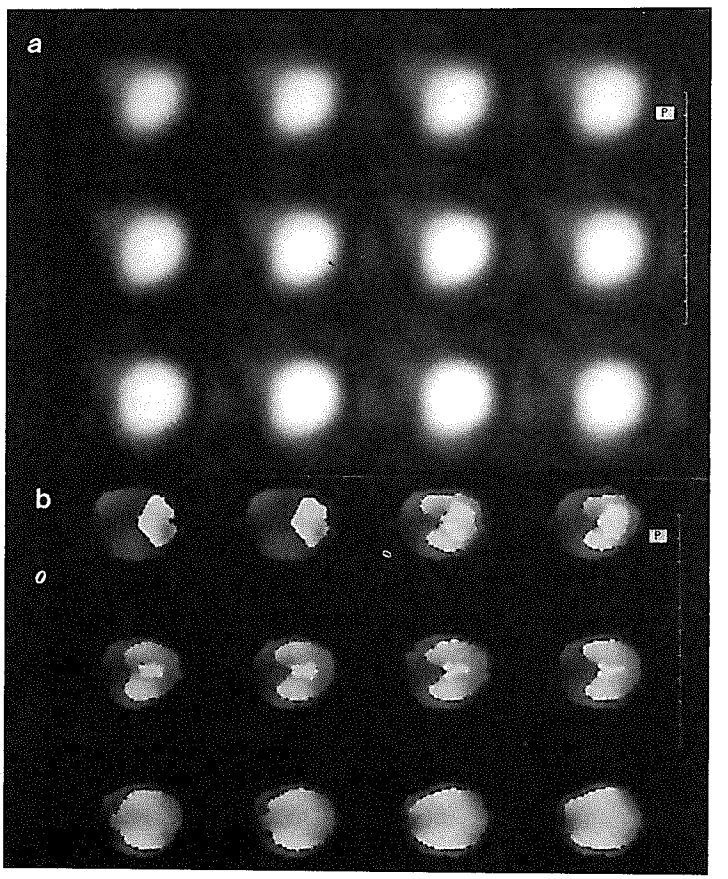

Figure 41. A portion of a 3-d PCr metabolite map reconstructed from the data set used in Fig. 40 with (a) the Fourier method and (b) the GSLIM method.

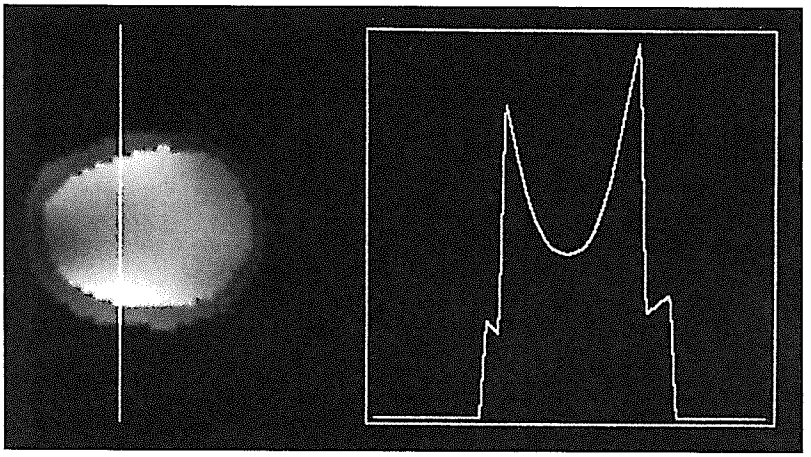

Figure 42. A single slice extracted from Fig. 45b to show the variations of the B_1 field and coil sensitivity revealed by the GSLIM reconstruction.

come by rapid processing capabilities of modern computer systems. Future developments will include better image models incorporating more realistic and powerful constraints and better computational procedures for parameter estimation.

In summary, constrained methods, when used properly, are extremely valuable for image reconstruction for limited data. We hope that this review would provide the reader a better understanding of many of the existing methods and encourage more research into this important area so that new reconstruction methods will continue to arise to meet the need of faster and more reliable and accurate imaging experiments.

Acknowledgments — This work is supported in part by NSF Science and Technology Grant STC-8920133, NIH Research Resource Grant 1P41PR05964-01, the National Center for Supercomputing Applications (NCSA), the Servants United Foundation, the Whitaker Foundation, and the National Sciences and Engineering Research Council (NSERC) of Canada.

This review would not have been possible without the support of the Biomedical Magnetic Resonance Laboratory (BMRL) of the University of Illinois at Urbana-Champaign; special thanks are due to the following members (past and present) of this laboratory for their helpful discussions and constructive criticisms of several methods reviewed in the paper: Prof. M. Joan Dawson; Clinton S. Potter; Dr. Carl Gregory; Dr. Haakil Lee (now at Vanderbilt University); Dr. Xiaohong Zhou (now at Duke University); to the following members of the Radiology Department of the University Hospitals of Cleveland for their helpful comments, contributions of data sets and figures, and photographic support: Sundar Amartur; Piotr Wielopolski; Jim Goldfarb; Joe Molter; Prof. Steve H. Izen; and to Dr. Andrain Crawley of the Department of Radiology of The Foothills Hospital, and Prof. Spence T. Nichols at the Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada, for their proofreading of the manuscript and constructive criticisms; and to Dr. Keith Thulborn of the Department of Radiology of Massachusetts General Hospital for commenting on the images.

The first author, Z.-P. Liang, would also like to extend his thanks to Prof. John C. Gore of Yale University and Prof. Mark Henkelman of the University of Toronto for their initial encouragement to write this review; to Ms. Emily C. Martin of Wolfram Research, Inc. and an anonymous referee for their constructive comments.

REFERENCES

1. J. B. Abbiss, M. Defrise, C. D. Mol, and H. S. Dhadwal. Regularized iterative and noniterative procedures for object restoration in the presence of noise: An error analysis. *J. Opt. Soc. Am. A.* 73:1470–1475; 1983.
2. J. W. Adams. A new optimal window. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-39:1753–1769; 1991.

3. D. F. Aldridge, S. E. Dosso, A. L. Endre, and D. W. Oldenburg. New methods for constructing flattest and smoothest models. *Inverse Problems* 7:499–513; 1991.
4. S. Amartur and E. M. Haacke. Modified iterative model based on data extrapolation method to reduce Gibbs ringing. *J. Magn. Res. Imaging* 1:307–317; 1991.
5. S. Amartur, Z.-P. Liang, F. Boada, and E. M. Haacke. A phase constrained data extrapolation method to reduce Gibbs ringing. *J. Magn. Reson. Imaging* 1:721–724; 1991.
6. H. C. Andrew and B. R. Hunt. *Digital Image Restoration*. Englewood Cliffs, NJ Prentice Hall; 1977.
7. K. S. Arun and L. C. Potter. Existence and uniqueness of band-limited positive semi-definite extrapolation. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-38:547–549, 1990.
8. W. P. Aue. Localization methods for *in vivo* nuclear magnetic resonance spectroscopy. *Rev. Magn. Reson. Med.* 1:21–72; 1986.
9. L. Auslander and F. A. Grunbaum. The Fourier transform and the discrete Fourier transform. *Inverse Problems* 5:149–164; 1989.
10. L. D. Bacon and R. E. Post. A limitation of the Kumaresan-Prony algorithm in direction-finding applications. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-32:912–914; 1984.
11. H. Barkhuijsen, R. D. Beer, W. M. M. J. Bovee, J. H. N. Creyghton, and D. Van Ormondt. Application of linear prediction and singular value decomposition (LPSVD) to determine NMR frequencies and intensities from the FID. *J. Magn. Reson.* 2:86–89; 1985.
12. H. Barkhuijsen, R. D. Beer, and D. V. Ormondt. Aspects of the computational efficiency of LPSVD. *J. Magn. Reson.* 64:343–346; 1985.
13. H. Barkhuijsen, R. D. Beer, and D. V. Ormondt. Error theory for time-domain signal analysis with linear prediction and singular value decomposition. *J. Magn. Reson.* 67:371–375; 1986.
14. H. Barkhuijsen, R. D. Beer, and D. V. Ormondt. Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals. *J. Magn. Reson.* 73:553–557; 1987.
15. C. W. Barnes. Object restoration in a diffraction-limited imaging system. *J. Opt. Soc. Am. A* 56:575–578; 1966.
16. I. Barrowdale and R. E. Erickson. Algorithms for least squares prediction and maximum entropy spectral analysis — Part I: Theory. *Geophysics* 43:420–432; 1980.
17. I. Barrowdale and R. E. Erickson. Algorithms for least squares prediction and maximum entropy spectral analysis—Part II: Fortran Program. *Geophysics* 43:433–446; 1980.
18. E. M. Bellon, E. M. Haacke, P. E. Coleman, D. C. Sacco, D. A. Steiger, R. E. Gangarosa. MR artifacts: A review. *AJR* 147:1271–1281; 1986.
19. M. Bendinelli, A. Consortini, L. Ronchi, and B. R. Frieden. Degrees of freedom, and eigenfunctions, for the noisy image. *J. Opt. Soc. Am. A* 64:1498–1502; 1974.
20. A. Ben-Israel. On error bounds for generalized inverses. *SIAM J. Numer. Anal.* 3:585–592; 1966.
21. N. Benvenuto and G. L. Pierobon. Extension of the sampling theorem to exponentially decaying causal functions. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP–39:189–190; 1991.
22. M. A. Bernstein, D. M. Thomasson, and W. H. Perman. Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase corrected real construction. *Med. Phys.* 16:813–817; 1989.
23. M. Bertero and E. R. Pike. Resolution in diffraction-limited imaging, a singular value analysis. *Optica Acta* 29:727–746; 1982.
24. M. Bertero, P. Brianzi, and E. R. Pike. On the recovery and resolution of exponential relaxation rates from experimental data: Laplace transform inversions in weighted spaces. *Inverse problems* 1:1–15; 1985.
25. M. Bertero, C. D. Mol, and E. R. Pike. Linear inversion with discrete data. I: General formulation and singular system analysis. *Inverse Problems* 1:301–330; 1985.
26. M. Bertero, C. D. Mol, and E. R. Pike. Linear inversion with discrete data. II: Stability and regularization. *Inverse Problems*. 4:573–594; 1988.
27. R. A. Blahut. *Fast Algorithms for Digital Signal Processing*. Reading, MA: Addition-Wesley; 1985.
28. F. Boada and E. M. Haacke. Effects of noise, phase and model order on the performance of the constrained reconstruction approach for magnetic resonance image reconstruction. J. Magn. Reson. (submitted); 1992.
29. A. W. Bojanczyk, R. P. Brent, and F. R. de Hoog. QR factorization of Toeplitz matrices. *Num. Math.* 49:18–94; 1986.
30. G. E. Box and G. M. Jenkins. *Times Series Analysis Forecasting and Control*. San Francisco, CA: Holden-Day; 1970.
31. R. N. Bracewell. *The Fourier Transform and its Applications*. New York: McGraw-Hill; 1986.

32. Y. Bresler and A. Macovski. Three-dimensional reconstruction from projections with incomplete and noisy data by object estimation. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-35:1139–1152; 1986.
33. E. O. Bringham. *The Fast Fourier Transform and its Applications*. Englewood, NJ: Prentice Hall; 1988.
34. M. J. Bronskill, E. R. McVeigh, W. Kucharczyk, and R. M. Henkelman. Syrinx-like artifacts on MR images of the spinal cord. *Radiology* 166:485–488; 1988.
35. H. R. Brooker, T. H. Mareci, and J. Mao. Selective Fourier transform localization. *Magn. Reson. Med.* 5:417–433; 1987.
36. T. R. Brown, B. M. Kincaid, and K. Ugurbil. NMR chemical shift imaging in three dimensions. *Proc. Natl. Acad. Sci. USA* 79:3523–3526; 1982.
37. J. R. Bunch. Stability of methods for solving Toeplitz systems of equations. *SIAM J. Sci. Stat. Comput.* 6:349–364; 1985.
38. S. F. Burch, S. F. Gull, and J. Skilling. Image restoration by a powerful maximum entropy method. *Computer Vision, Graphics, and Image Processing* 23:113–128, 1983.
39. J. P. Burg. *Maximum Entropy Spectrum Analysis*, Doctoral dissertation, Stanford University, Palo Alto, CA; 1975.
40. J. A. Cadzow. An extrapolation procedure for band-limited signals. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-27:4–12; 1979.
41. P. T. Callaghan and C. D. Eccles. Sensitivity and resolution in NMR imaging. *J. Magn. Reson.* 71:426–445; 1987.
42. D. C. Champency. *A Handbook of Fourier Theorems*. Cambridge, MA: Cambridge University Press; 1987.
43. Z. H. Cho, H. S. Kim, H. B. Song, and J. Cumming. Fourier transform nuclear magnetic resonance tomographic imaging. *Proc. IEEE* 70:1152–1173; 1982.
44. R. T. Constable and R. M. Henkelman. Contrast, resolution, and detectability. *J. Comp. Asst. Tomogr.* 15:297–303; 1991.
45. R. T. Constable and R. M. Henkelman. Data extrapolation for truncation artifact removal. *Magn. Reson. Med.* 17:108–118, 1991.
46. R. T. Constable and R. M. Henkelman. Maximum entropy reconstruction in MRI. In *Maximum Entropy and Bayesian Methods*, J. Skilling (ed.), New York: Kluwar: 175–180; 1989.
47. R. T. Constable and R. M. Henkelman. Why MEM does not work in MR image reconstruction. *Magn. Reson. Med.* 14:12–25; 1990.
48. J. J. Cuppen and A. Van Est. Reducing MR imaging time by one-sided reconstruction. *Topical Conf. Fast MRI Techniques*, Cleveland, Ohio 1987.
49. G. Cybenko. Locations of zeros of predictor polynomials. *IEEE Trans. Automat. Contr.* AC-27:235–237; 1982.
50. G. Cybenko. Fast Toeplitz orthogonalization using inner products. *SIAM J. Sci. Stat. Comput.* 8:734–740; 1987.
51. I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* 41:909–996; 1988.
52. M. E. Davison. The ill-conditioned nature of the limited angle tomography problem. *SIAM J. Appl. Math.* 43:428–448; 1983.
53. A. A. DeGraaf and W. M. M. J. Bovee. Improved quantification of in vivo ^{1}H NMR spectra by optimization of signal acquisition and processing and by incorporation of prior knowledge into the spectral fitting. *Magn. Reson. Med.* 15:305–319; 1990.
54. M. A. Delsuc, F. Ni, and G. C. Levy. Improvement of linear prediction processing of NMR spectra having very low signal-to-noise. *J. Magn. Reson.* 73:548–552; 1987.
55. J. E. Dennis, Jr. A user's guide to nonlinear optimation algorithms. *Proc. IEEE* 72:1765–1776; 1984.
56. B. C. De Simone, F. De Luca, and B. Maraviglia. Maximum entropy method in phase-encoding NMR imaging. *Magn. Reson. Med.* 4:78–82; 1987.
57. B. C. De Simone, F. De Luca, and B. Maraviglia. Maximum entropy and 2DFT NMR images. *Magn. Reson. Med.* 8:332–339; 1988.
58. K. Derby, C. Hawryszko, and J. Tropp. Baseline deconvolution, phase correction, and signal quantification in Fourier localized spectroscopic imaging. *Magn. Reson. Med.* 12:235–240; 1989.
59. E. J. Diethorn and D. C. Munson, Jr. A linear, time-varying system framework for noniterative discrete-time band-limited signal extrapolation. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-39:55–68; 1992.

60. G. T. di Francia. Resolving power and information. *J. Opt. Soc. Am. A* 45:497–501; 1955.
61. D. L. Donoho, I. M. Johnstone, A. S. Stern, and J. C. Hoch. Does the maximum entropy method improve sensitivity. *Proc. Natl. Acad. Sci. USA* 87:5066–5068; 1990.
62. W. A. Edelstein, G. H. Glover, C. J. Hardy, and K. W. Redington. The intrinsic signal-to-noise ratio in NMR imaging. *Magn. Reson. Med.* 3:604–618; 1986.
63. J. Eriksson. A note on solution of large sparse maximum entropy problems with linear equality constraints. *Math. Programming* 18:146–154; 1980.
64. D. A. Feinberg, J. D. Hale, J. C. Watts, L. Kaufman, and A. Mark. Halving MR imaging time by conjugation: Demonstration at 3.5 KG. *Radiology* 161:527–531; 1986.
65. R. Fletcher. *Practical Methods of Optimization*. New York: John Wiley & Sons; 1987.
66. M. Fuderer. Adding noise to improve image quality. *Book of Abstracts, 8th Ann. Meet. Soc. Magn. Reson. Med.* Amsterdam, Netherlands: 913; 1989.
67. R. W. Gerchberg. Superresolution through error energy reduction. *Optica Acta* 21:709–720; 1974.
68. D. A. Girard. Optimal regularized reconstruction in computerized tomography. *SIAM J. Sci. Stat. Comput.* 8:934–950; 1987.
69. F. Godtliebsen and E. Spjotvoll. Comparison of statistical methods in MR imaging. *Int. J. Imaging Syst. Techn.* 3:33–39; 1991.
70. F. Godtliebsen. Noise reduction using Markov random fields. *J. Magn. Reson.* 92:102–114; 1991.
71. J. W. Goldfarb, E. M. Haacke, F. Boada, and W. Lin. Reconstruction of doubly asymmetric data for 2D rapid imaging. *Book of Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Med.* San Francisco, CA: 1226; 1991.
72. G. H. Golub and V. Pereyra. The differentiation of pseudo-inverses and non-linear least squares problems whose variables separate. *SIAM J. Numer. Anal.* 10:413–432; 1973.
73. G. H. Golub and C. F. Van Loan. *Matrix Computations*, 2nd ed. Baltimore, MD: The Johns Hopkins University Press; 1989.
74. S. F. Gull and J. Daniel. Image reconstruction from incomplete and noisy data. *Nature* 272:686–690; 1978.
75. S. F. Gull and J. Skilling. Maximum entropy method in image processing. *Proc. IEE.* 131:646–659; 1984.
76. Q. Guo, G. Kashmar, and O. Nalcioglu. NMR angiography with enhanced quasi-half-echo scanning. *Magn. Reson. Imaging* 9:129–139; 1991.
77. I. Guttman, V. Pereyra, and H. D. Scolnik. Least squares estimation for a class of non-linear problems. *Technometrics* 15:209–218; 1973.
78. E. M. Haacke. The effects of finite sampling in spin-echo or field-echo magnetic resonance imaging. *Magn. Reson. Med.* 4:407–421; 1987.
79. E. M. Haacke. Solving for nonideal conditions in two-dimensional Fourier transform magnetic resonance imaging using a generalized inverse transform. *Inverse Problems* 3:421–435; 1987.
80. E. M. Haacke and Z.-P. Liang. High resolution, limited view reconstructions. *Topical Conf. Fast MRI Techniques*. Cleveland, OH; 1987.
81. E. M. Haacke, Z.-P. Liang, and S. H. Izen. Superresolution reconstruction through object modeling and estimation. *IEEE Trans. Acoust., Speech, Signal Processing* 37:592–595; 1989.
82. E. M. Haacke, Z.-P. Liang, and S. H. Izen. Constrained reconstruction: A super-resolution optimal signal-to-noise alternative to the Fourier transform. *Med. Phys.* 16:388–397; 1989.
83. E. M. Haacke, Z.-P. Liang, and F. Boada. Image reconstruction using projection onto convex sets, model constraints, and linear prediction theory for the removal of phase, motion and Gibbs artifacts in magnetic resonance and ultrasound imaging. *Opt. Eng.* 5:555–566; 1990.
84. E. M. Haacke, J. A. Tkach, and T. B. Parrish. Reduction of T_2^* dephasing in gradient field-echo imaging. *Radiology* 170:457–462, 1989.
85. E. M. Haacke, J. Mitchell, and D. Lee. Improved contrast at 1.5 Tesla using half Fourier imaging: Application to spin-echo and angiographic imaging. *Magn. Reson. Imaging* 8:79–90; 1990.
86. E. M. Haacke, E. D. Lindskog, and W. Lin. A fast, iterative, partial-Fourier technique capable of local phase recovery. *J. Magn. Reson.* 92:126–145; 1991.
87. K. M. Hanson and G. W. Wecksung. Local basis-function approach to computed tomography. *Appl. Opt.* 24:4028–4039; 1985.
88. V. J. Harding and B. A. Eagles. The creative content of brain. *J. Biol. Chem.* 60:301–310; 1924.
89. F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* 66:51–83; 1978.

90. J. L. Harris. Diffraction and resolving power. *J. Opt. Soc. Am.* 54:931–936; 1964.
91. J. L. Harris. Image evaluation and restoration. *J. Opt. Soc. Am.* 56:569–574; 1966.
92. D. A. Hayner and W. K. Jenkins. The missing cone problem in computer tomography. In *Advances in Computer Vision and Image Processing*, vol. 1, T. S. Huong, ed. Greenwich, CT, JAI Press Inc.: 83–144; 1984.
93. R. M. Henkelman. Measurement of signal intensities in the presence of noise in MR images. *Med. Phys.* 12:232–233; 1985.
94. R. M. Henkelman and M. J. Bronskill. Artifacts in magnetic resonance imaging. *Rev. Magn. Reson. Med.* 2:1–126; 1987.
95. E. A. Heredia. Modeling and order estimation of magnetic resonance images. Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, September, 1991.
96. E. A. Heredia, M. R. Smith, and S. T. Nichols. Order estimation in ARMA-based MRI reconstruction. *Book of Abstracts, 9th Ann. Meet. Soc. Magn. Reson. Imaging,* Chicago, IL: 159; 1991.
97. W. S. Hinshaw and A. H. Lent. An introduction to NMR imaging: From the Bloch equation to the imaging equation. *Proc. IEEE* 71:338–350; 1983.
98. X. Hu, D. N. Levin, P. C. Lauterbur, and T. Spraggins. SLIM: Spectral Localization by Imaging. *Magn. Reson. Med.* 8:314–322; 1988.
99. X. Hu, V. Johnson, W. H. Wong, and C.-T. Chien. Bayesian image processing in magnetic resonance imaging. *Magn. Reson. Imaging* 9:611–620; 1991.
100. T. S. Huang, J. L. C. Sanz, H. Fan, J. Shafii, and B.-M. Tsai. Numerical comparison of several algorithms for band-limited signal extrapolation. *Applied Optics* 23:307–317; 1984.
101. T. S. Huang. The subjective effect of two-dimensional pictorial noise. *IEEE Trans. Information Theory* IT-11:43–53; 1965.
102. J. Jackson, A. Macovski, and D. Nishimura. Low frequency restoration. *Magn. Reson. Med.* 11:248–257; 1989.
103. A. K. Jain and S. Ranganath. Extrapolation algorithms for discrete signals with application in spectral estimation. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-29:830–845; 1981.
104. V. K. Jain and T. K. Sarkar. High-performance signal-modeling by pencil-of-functions method: band-pass case. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-34:997–1000; 1986.
105. V. K. Jain, T. K. Sarkar, and D. D. Weiner. Rational modeling by pencil-of-functions method. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-31:564–573; 1983.
106. P. A. Jansson. *Deconvolution with applications in Spectroscopy*. New York: Academic Press; 1984.
107. E. T. Jaynes. On the rationale of maximum-entropy methods. *Proc. IEEE* 70:939–952; 1982.
108. E. T. Jaynes. Prior probabilities. *IEEE Trans. Syst. Sci. Cybern.* SSC-4:227–241; 1968.
109. M. A. Jenkins and J. F. Traub. A three-stage algorithm for real polynomials using quadratic iteration. *SIAM J. Numer. Anal.* 7:252–263; 1970.
110. M. Joilot, B. M. Mazoyer, and R. Huesman. In vivo NMR spectral estimation: A comparison between time and frequency domain methods. *Magn. Reson. Med.* 18:358–370; 1991.
111. J. A. Jones and P. J. Hore. The Maximum entropy method and Fourier transform compared. *J. Magn. Reson.* 92:276–292; 1991.
112. J. A. Jones and P. J. Hore. The Maximum entropy method. Appearance and reality. *J. Magn. Reson.* 92:363–376; 1991.
113. N. Kalouptsidis, G. Carrayaannis, and D. Manolakis. Fast algorithms for block Toeplitz matrices with Toeplitz entries. *Signal Processing* 6:77–81; 1984.
114. A. K. Katsaggelos and S. M. Estratiadis. A class of iterative signal restoration algorithms. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-38:778–786; 1990.
115. L. Kaufman and V. Pereyra. A method for separable nonlinear least squares problems with separable nonlinear equality constraints. *SIAM J. Numer. Anal.* 15:12–20; 1978.
116. S. Kawata, K. Minami, and S. Minami. Superresolution of Fourier transform spectroscopy data by the maximum entropy method. *Applied Optics* 22:3593–3598; 1983.
117. S. M. Kay and S. L. Marple, Jr. Spectrum analysis—A modern perspective. *Proc. IEEE* 69:1380–1419; 1981.
118. P. Knapp. The determination of MA model parameters. Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, September, 1986.
119. F. N. Kong. A General framework for developing and evaluating band-limited function extrapolation methods. *Proc. IEEE* 76:1381–1383; 1988.

120. S. Kopec. Properties of maximum entropy approximate solutions to Fredholm integral equations. *J. Math. Phys.* 32:1269–1272; 1991.
121. E. L. Kosarev. Shannon's superresolution limit for signal recovery. *Inverse Problems* 6:55–76; 1990.
122. H. Krishna and S. D. Morgera. The Levinson recurrence and fast algorithms for solving Toeplitz systems of linear equations. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-35:839–848; 1984.
123. A. Kumar, D. Welti, and R. Ernst. NMR Fourier zeugmatography. *J. Magn. Reson.* 18:69–83; 1975.
124. R. Kumaresan and D. W. Tufts. Estimating the parameters of exponentially damped sinusoid and pole-zero modeling in noise. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-30:217–220; 1982.
125. R. Kumaresan, L. L. Scharf, and A. K. Shaw. An Algorithm for pole-zero modeling and spectral analysis. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-34:637–640; 1986.
126. S. Y. Kung, K. S. Arun, and B. D. Rao. State-space singular value decomposition based methods for the harmonic retrieval problem. *J. Opt. Soc. Am. A* 73:1799–1811; 1983.
127. A. Lannes and S. Roques. Resolution and robustness in image processing: A new regularization principle. *J. Opt. Soc. Am. A* 4:189–199; 1987.
128. P. C. Lauterbur. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature* 242:190–191; 1973.
129. P. C. Lauterbur, D. M. Kramer, W. V. House, Jr., C.-N. Chen. Zeugmatographic high resolution nuclear magnetic resonance spectroscopy: Images of chemical inhomogeneity within macroscopic objects. *J. Am. Chem. Soc.* 97:6866–6868; 1975.
130. P. C. Lauterbur, D. N. Levin, and R. B. Marr. Theory and simulation of NMR spectroscopic imaging and field plotting by projection reconstruction involving an intrinsic frequency dimension. *J. Magn. Reson.* 59:536–541; 1984.
131. C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall; 1987.
132. J.-S. Lee. Image smoothing and the sigma filter. *Computer Vision, Graphics and Image Processing,* 24:255–269; 1983.
133. A. Lent. An iterative method for the extrapolation of band-limited functions. *J. Math. Anal. Appl.* 83:554–565; 1981.
134. Z.-P. Liang. *Constrained Reconstruction from Incomplete and Noisy Data: A New Parametric Approach,* Doctoral dissertation, Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, May, 1989.
135. Z.-P. Liang, E. M. Haacke, and C. W. Thomas. High resolution spectral estimation through a localized polynomial approximation. *Inverse Problems* 5:831–847; 1989.
136. Z.-P. Liang and P. C. Lauterbur. A generalized series approach to MR spectroscopic imaging. *IEEE Trans. Med. Imaging.* MI-10:132–137; 1991.
137. Z.-P. Liang and E. M. Haacke. Parametric imaging: Overcoming the Gibbs artifact. *Abstracts, 8th Ann. Meet. Soc. Magn. Reson. Imaging,* Washington, DC: 75; 1990.
138. Z.-P. Liang. Optimal half-Fourier reconstruction: A generalized series approach. *Book of Abstracts, 9th Ann. Meet. Soc. Magn. Reson. Med.* New York: 552; 1990.
139. Z.-P. Liang, H. Lee, and P. C. Lauterbur. On errors of the SLIM technique. *Book of Abstracts, 9th Ann. Meet. Soc. Magn. Reson. Med.* New York: 1077;1990.
140. Z.-P. Liang and P. C. Lauterbur. A theoretical analysis of the SLIM technique. *J. Magn. Reson.* (submitted); 1992.
141. Z.-P. Liang and P. C. Lauterbur. Effective use of anatomical information for spectroscopic imaging by GSLIM. *Works-in-Progress Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Med.* San Francisco, CA: 1004; 1991.
142. Z.-P. Liang, C. Gregory, M. J. Dawson, and P. C. Lauterbur. An improved linear prediction method for correction of phase and baseline artifacts. *Works-in-Progress Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Med.* San Francisco, CA: 1169; 1991.
143. Z.-P. Liang, C. Gregory, M. J. Dawson, and P. C. Lauterbur. Utilizing the data in the phase-encoding "dead-time" by modified SLIM processing. *Works-in-Progress Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Med.* San Francisco, CA: 168; 1991.
144. Z.-P. Liang and P. C. Lauterbur. Improved temporal/spatial resolution in functional imaging through-generalized series reconstruction. *Works-in-Progress Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Imaging,* New York: s15; 1992.
145. Z.-P. Liang and P. C. Lauterbur. A generalized series reconstruction algorithm for fast time-sequential magnetic resonance imaging. *IEEE Trans. Med. Imaging* (submitted; 1992).

146. D. M. Lin and E. K. Wong. A survey on the maximum entropy method and parameter estimation. *Physics Reports* 193:41–135; 1990.
147. E. D. Lindskog. *Partial Fourier Imaging,* Master of Science Thesis, Department of Electrical Engineering and Applied Physics, Case Western Reserve University, Cleveland, 1989.
148. R. B. Lufkin, E. Pusey, D. D. Stark, R. Brown, B. Leikind, and W. N. Hanafee. Boundary artifact due to truncation errors in MR imaging. *AJR* 147:1283–1287; 1986.
149. A. Macovski and D. Spielman. *In vivo* spectroscopic magnetic resonance imaging using estimation theory. *Magn. Reson. Med.* 3:97–104; 1986.
150. J. Makhoul. On the eigenvectors of symmetric Toeplitz matrices. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-29:868–872; 1981.
151. J. Makhoul. Spectral inear prediction: Properties and applications. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-32:283–296; 1974.
152. J. Makhoul. Linear prediction: A tutorial review. *Proc. IEEE* 63:561–580; 1975.
153. S. G. Mallat. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Machine Intell.* 11:674–693; 1989.
154. R. J. Mammone. Spectral extrapolation of constrained signals. *J. Opt. Soc. Am. A* 73:1476–1480; 1983.
155. T. H. Mareci and H. R. Brooker. Essential considerations for spectral localization using indirect gradient encoding of spatial information. *J. Magn. Reson.* 92:229–246;1991.
156. P. Margosian, F. Schmitt, and D. E. Purdy. Faster MR imaging: Imaging with half the data. *Health Care Instr.* 1:195–197; 1986.
157. S. L. Marple, Jr. *Digital Spectral Analysis with Applications.* Englewood Cliffs, NJ: Prentice-Hall; 1987.
158. S. L. Marple, Jr. Frequency resolution of Fourier and maximum entropy spectral estimates. *Geophysics* 47:1303–1307; 1982.
159. J. F. Martin and C. F. Tirendi. Modified linear prediction modeling in magnetic resonance imaging. *J. Magn. Reson.* 82:392–399; 1989.
160. J. F. Martin. The maximum entropy method in NMR. *J. Magn. Reson.* 65:291–297; 1985.
161. S. R. McFall, N. J. Pelc, and R. M. Vaurek. Correction of spatially dependent phase shifts for partial Fourier Imaging. *Magn. Reson. Imaging* 6:143–155; 1988.
162. G. McGibney. *Phase Sensitive Reconstruction of Magnetic Resonance Images,* Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, September, 1991.
163. G. McGibney, M. R. Smith and S. T. Nichols. Quantitative evaluation of several phase estimation algorithms used in MRI. *Magn. Reson. Med.* 1992 (in press).
164. G. McGibney, M. R. Smith, and S. T. Nichols. Estimation of phase errors in magnetic resonance imaging. *Book of Abstracts, 9th Ann. Meet. Soc. Magn. Reson. Med.* New York: 563; 1990.
165. D. K. Mitchell. *Image Enhancement in Magnetic Resonance Imaging,* Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, June, 1987.
166. D. K. Mitchell, S. T. Nichols, M. R. Smith, and K. Scott. The use of band selectable digital filtering in magnetic resonance image enhancement. *Magn. Reson. Med.* 9:353–368; 1989.
167. D. K. Mitchell, M. R. Smith, and S. T. Nichols. Efficient zooming (interpolation) of MRI images. *Book of Abstracts, 6th Ann. Meet. Soc. Magn. Reson. Med.* New York: 765; 1987.
168. J. D. Mitchell. *Half-Fourier Imaging in Nuclear Magnetic Resonance,* Master of Science Thesis, Department of Biomedical Engineering, Case Western Reserve University, Cleveland, 1988.
169. P. R. Moran. Observations on maximum entropy processing of MR imaging. *Magn. Reson. Imaging.* 9:213–222; 1991.
170. J. P. Mugler III. Potential degradation in image quality due to selective averaging of phase-encoding line in Fourier transform MRI. *Magn. Reson. Med.* 19:170–174; 1991.
171. J. P. Mugler III and J. R. Brookeman. Evaluation of a simple method for reconstructing asymmetrically sampled echo data. *J. Magn. Res. Imaging* 1:487–491; 1991.
172. S. W. Nichols. The design of a hardware processor capable of performing autoregressive modeling in real time, Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, October, 1986.
173. D. C. Noll, D. G. Nishimura, A. Macovski. Homodyne detection in magnetic resonance imaging. *IEEE Trans. Med. Imaging* MI-10: 154–163; 1991.
174. A. Papoulis and M. S. Bertran. Digital filtering and prolate functions. *IEEE Trans. Circuit Theory* CT-19: 674–681; 1972.
175. A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Trans. Circuits Syst.* CAS-22: 735–742; 1975.

176. A. Papoulis. *Signal Analysis*. New York: McGraw–Hill; 1977.
177. A. Papoulis. Maximum entropy and spectral estimation: A review. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-29: 1176–1186; 1981.
178. D. L. Parker, G. T. Gullberg, and P. R. Frederick. Gibbs artifact removal in magnetic resonance imaging. *Med. Phys.* 14:640–645; 1987.
179. D. L. Parker and G. T. Gullberg. Signal-to-noise efficiency in magnetic resonance imaging. *Med. Phys.* 17:250–257; 1990.
180. A. R. Penner, M. R. Smith, and S. T. Nichols. Noise reduction in magnetic resonance images using DFT and TERA modeling reconstruction. *Proc. 11th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.:* 642–643; 1989.
181. E. R. Pike, J. G. McWhirter, M. Bertero, and C. de Mol. Generalised information theory for inverse problems in signal processing. *Proc. IEEE* 131:660–667; 1984.
182. W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling. *Numerical Recipes*. Cambridge, MA: Cambridge University Press; 1986.
183. M. A. Rahman and K. B. Yu. Total least squares approach for frequency estimation using linear prediction. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-35:1440–1454; 1987.
184. R. Rangayyan, A. P. Dhawan, and R. Gordon. Algorithm for limited-view computed tomography: An annotated bibliography and a challenge. *Applied Optics* 24:4000–4012; 1985.
185. B. D. Rao and K. S. Arun. Model based processing of signals: A state space approach. *Proc. IEEE* 80:283–309; 1992.
186. R. Ray, B. G. Sumpter, G. A. Pfeffer, S. K. Gray, and D. W. Noid. Novel methods for spectral analysis. *Physics Reports* 205:109–152; 1991.
187. S. J. Riederer. Recent advances in magnetic resonance imaging. *Proc. IEEE* 76:1095–1105; 1988.
188. D. J. Rossi and A. S. Willsky. Reconstruction from projections based on detection and estimation of objects — Part I and II: Performance analysis and robustness analysis. *IEEE Trans. Acoust., Speech, Signal Processing*. ASSP-32:886–906; 1984.
189. O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE SP Magazine:* 14–38: Oct. 1991.
190. V. M. Runge and M. L. Wood. Half-Fourier MR imaging of CNS disease. *Am. J. Neuroradiol.* 11:77–82; 1990.
191. C. K. Rushforth. Restoration, resolution, and noise. *J. Opt. Soc. Am. A* 58:539–545; 1968.
192. W. M. Saar. *High Quality Magnetic Resonance Imaging Using Reduced Data Sets and ARMA Modeling*, Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, November, 1988.
193. M. S. Sabri and W. Steenaart. An approach to band-limited signal extrapolation: The extrapolation matrix. *IEEE Trans. Circuits Sys.* CAS-25:74–78; 1978.
194. P. Sain-Marc, J.-S. Chen, and G. Medioni. Adaptive smoothing: A general tool for early vision. *IEEE Trans. Pattern Anal. Machine Intell.* 13:514–529; 1991.
195. J. L. C. Sanz and T. S. Huang. Iterative time-limited signal restoration. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-31:643–649; 1983.
196. J. L. C. Sanz and T. S. Huang. Discrete and continuous band-limited signal extrapolation. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-31:1276–1285; 1983.
197. J. L. C. Sanz and T. S. Huang. Unified Hilbert space approach to iterative least-squares linear signal restoration. *J. Opt. Soc. Am. A* 73:1455–1465; 1983.
198. J. L. C. Sanz and T. S. Huang. Some aspects of band-limited signal extrapolation: Models, discrete approximation, and noise. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-31:1492–1501; 1983.
199. J. L. C. Sanz and T. S. Huang. On the Gerchberg-Papoulis algorithm. *IEEE Trans. Circuits Syst.* CAS-30:907–908; 1983.
200. J. L. C. Sanz and T. S. Huang. Continuation techniques for a certain class of analytic functions. *SIAM J. Appl. Math.* 44:819–838; 1984.
201. T. K. Sarkar, D. D. Weiner, and V. K. Jain. Some mathematical considerations in dealing with the inverse problem. *IEEE Trans. Antennas Propagat.* AP-29:373–379; 1981.
202. T. K. Sarkar, K. R. Siarkiewicz, and R. F. Stratton. Survey of numerical methods for solution of large systems of linear equations for electromagnetic field problems. *IEEE Trans. Antennas Propagat.* AP-29: 847–856; 1981.
203. R. W. Schafer, R. M. Mersereau, and M. A. Richards. Constrained iterative restoration algorithms. *Proc. IEEE* 69:432–450, 1981.

204. T. K. Sarkar, D. D. Weiner, J. Nebat, and V. K. Jain. A discussion of various approaches to the identification/approximation problem. *IEEE Trans. Antennas Propagat.* AP-30:373–379; 1982.
205. A. E. Schussheim and D. Cowburn. Deconvolution of high-resolution two-dimensional NMR signals by digital signal processing with linear predictive singular value decomposition. *J. Magn. Reson.* 71:371–378; 1987.
206. K. Scott, D. Mitchell, S. T. Nichols, and M. R. Smith. Modeling of MRI subimages — A computationally efficient approach. *Book of Abstracts, 6th Ann. Meet. Soc. Magn. Reson. Med.* New York: 268; 1987.
207. M. I. Sezan and H. Stark. Image restoration by the method of convex projection: Part 2 — Applications and numerical results. *IEEE Trans. Med. Imaging* MI-1:95–101; 1982.
208. M. I. Sezan and H. Stark. Tomographic image reconstruction from incomplete view data by convex projections and direct Fourier inversion. *IEEE Trans. Med. Imaging,* MI-3:91–98, 1984.
209. J. E. Shore and R. W. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inform. Theory* 26:26–37; 1980.
210. S. Singh, S. N. Tandon, and H. M. Gupta. An iterative restoration technique. *Signal Processing* 11:1–11; 1986.
211. D. Slepian and H. O. Pollack. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. *Bell Syst. Tech. J.* 40:43–60; 1961.
212. D. Slepian. On bandwidth. *Proc. IEEE* 64:292–300; 1976.
213. T. J. Smit. Design and implementation of a high speed digital signal processor. Master of Science Thesis, Department of Electrical and Computer Engineering, University of Calgary, Canada, February, 1989.
214. M. R. Smith, S. T. Nichols, R. M. Henkelman, and M. L. Wood. Application of autoregressive modeling in magnetic resonance imaging to remove noise and truncation artifacts. *Magn. Reson. Imaging* 4:257–261; 1986.
215. M. R. Smith, W. Saar, and S. T. Nichols. High quality MR imaging with decreased acquisition time using autoregressive moving average modeling. *Works-in-Progress Abstracts, 5th Ann. Meet. Soc. Magn. Reson. Med.* Montreal, Quebec; 1986:81–82.
216. M. R. Smith, S. T. Nichols, R. M. Henkelman, and M. L. Wood. Application of auto-regressive moving average parametric modeling in magnetic resonance image reconstruction. *IEEE Trans. Med. Imaging* MI-5:132–139; 1986.
217. M. R. Smith and S. T. Nichols. Efficient algorithms for generating interpolated (zoomed) MR images. *Magn. Reson. Med.* 7:156–171; 1988.
218. M. R. Smith and S. T. Nichols. A two-dimensional modeling reconstruction technique for magnetic resonance data. *J. Magn. Reson.* 85:573–580; 1989.
219. M. R. Smith and S. T. Nichols. A comparison of models used as alternative magnetic resonance image reconstruction methods. *Magn. Reson. Imaging* 8:173–183; 1990.
220. M. R. Smith, S. T. Nichols, R. T. Constable, and R. M. Henkelman. A quantitative comparison of the TERA modeling and the DFT magnetic resonance image reconstruction technique. *Magn. Reson. Med.* 19:1–19; 1991.
221. M. R. Smith, S. T. Nichols, and A. Crawley. The subjective effect of noise on MRI images reconstructed from short data sets. *IEEE Trans. Med. Imaging* December (submitted); 1991.
222. M. R. Smith, G. McGibney, and S. T. Nichols. Quantitative evaluation of a combined partial Fourier and modeling algorithm for use in MRI reconstruction. *Magn. Reson. Med.* (submitted); 1992.
223. M. R. Smith and S. T. Nichols. DFT matching — A method for adding constraints to alternative MRI Reconstruction Algorithms. *Book of Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Med.* 2:749; 1991.
224. M. R. Smith, S. T. Nichols, R. T. Constable, and R. M. Henkelman. A quantitative comparison of TERA modeling and the DFT magnetic resonance image reconstruction techniques. *Magn. Reson. Med.* 19:1–19; 1991.
225. H. V. Sorensen, D. L. Jones, M. T. Heideman, C. S. Burrus. Real-valued fast Fourier transform algorithms. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-35:849–863; 2987.
226. H. Stark. *Image Recovery: Theory and Application.* New York: Academic Press; 1987.
227. A. S. Stern and J. C. Hoch. A new, storage-efficient algorithm for maximum-entropy spectrum reconstruction. *J. Magn. Reson.* 97:255–270; 1992.
228. G. W. Stewart. On the perturbation of pseudo-inverses, projections and linear least squares problems. *SIAM Review* 19:634–662; 1977.
229. A. E. Stillman, D. N. Levin, D. B. Yang, R. B. Marr, and P. C. Lauterbur. Back projection reconstruction of spectroscopic NMR images from incomplete sets of projections. *J. Magn. Reson.* 69:168–175; 1986.

230. D. R. Sweet. Fast Toeplitz orthogonalization. *Num. Math.* 43:1–21; 1984.
231. K. Tanabe. Projection method for solving a singular system of linear equations and its applications. *Num. Math.* 17:203–214; 1971.
232. C. F. Tirendi and J. F. Martin. Quantitative analysis of NMR spectra by linear prediction and total least squares. *J. Magn. Reson.* 85:162–169; 1989.
233. H. J. Trussell. Convergence criteria for iterative restoration methods. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-31:129–136; 1983.
234. H. J. Trussell and M. R. Civanlar. The Landweber iteration and projection onto convex sets. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-33: 1632–1633; 1985.
235. H. J. Trussell and M. R. Civanlar. The feasible solution in signal restoration. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-32:201–212; 1984.
236. D. W. Tufts, R. Kumaresan, and I. Kirsteins. Data adaptive signal estimation by singular value decomposition of a data matrix. *Proc. IEEE* 70:684–685; 1982.
237. D. A. Turner, M. I. Rapoport, W. D. Erwin, M. McGould, and R. I. Silvers. Truncation artifact: A potential pitfall in MR imaging of the menisci of the knee. *Radiology* 179:629–633; 1991.
238. D. B. Twieg. The *k*-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. *Med. Phys.* 10:610–621; 1983.
239. J. W. C. Van Der Veen, R. De Beer, P. R. Luyten, and D. Van Ormondt. Accurate Quantification of in vivo ^{31}P NMR signals using the variable projection method and prior knowledge. *Magn. Reson. Med.* 6:92–98; 1988.
240. V. Viti, P. Barone, L. Guidoni, and E. Massaro. Maximum entropy spectral analysis of ^{31}P NMR signals from human cells. *J. Magn. Reson.* 67:91–102; 1986.
241. V. Viti, E. Massaro, L. Guidoni, and P. Barone. The use of the maximum entropy method in NMR spectroscopy. *J. Magn. Reson.* 70:379–393; 1986.
242. M. von Kienlin and R. Mejia. Spectral localization with optimal pointspread function. *J. Magn. Reson.* 94:268–287; 1991.
243. M. Wax and T. Kailath. Efficient inversion of Toeplitz-block Toeplitz matrix. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-31:1218–1221; 1983.
244. J. B. Weaver, D. M. Healy, Y. Xu, and L. D. Cromwell. Optimum temporal filters and multiresolution representations of MR images. *Book of Abstracts, 9th Ann. Meet. Soc. Magn. Reson. Med.* New York: 567; 1990.
245. J. B. Weaver, Y. Xu, D. M. Healy, and L. D. Cromwell. Filtering noise images with the Wavelet transform. *Magn. Reson. Med.* 21:288–295; 1991.
246. J. B. Weaver, Y. Xu, D. M. Healy, and J. R. Driscoll. Wavelet-encoded MR imaging. *Magn. Reson. Med.* 24:275–287; 1992.
247. A. Webb, Z.-P. Liang, R. L. Magin, and P. C. Lauterbur. Application of a generalized series reconstruction algorithm to biological MRI. *Works-in-Progress Abstracts, 10th Ann. Meet. Soc. Magn. Reson. Imaging* New York: s26; 1992.
248. M. L. Wood and R. M. Henkelman. Truncation artifacts in magnetic resonance imaging. *Magn. Reson. Med.* 2:517–526; 1985.
249. H. Yan and J. C. Gore. The relation of HSVD to LPSVD for fitting time domain signals. *J. Magn. Reson.* 80:324–327; 1988.
250. H. Yan, J. C. Gore. An efficient algorithm for MR image reconstruction without low spatial frequencies. *IEEE Trans. Med. Imaging* MI-9:184–189; 1990.
251. H. Yan and J. Mao. The relation of low frequency restoration methods to the Gerchberg-Papoulis algorithm. *Magn. Reson. Med.* 16:166–172; 1990.
252. H. Yan and J. C. Gore. Improved parametric spectral estimation using a peak extraction technique. *J. Magn. Reson.* 81:400–405; 1989.
253. H. Yan and M. Braun. Image reconstruction from Fourier domain data sampled along a zig-zag trajectory. *Magn. Reson. Med.* 18:405–410; 1991.
254. D. C. Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Trans. Circuits Sys.* CAS-25:694–702; 1978.
255. D. C. Youla and H. Webb. Image restoration by the method of convex projections: Part I-Theory. *IEEE Trans. Med. Imaging* MI-1:81–94; 1982.
256. G. U. Yule. On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Phil. Trans. R. Soc. London, ser. A.* 226:267–298; 1927.

257. L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai. On detection of the number of signals in presence of white noise. *J. Multivariate Anal.* 20:1–25; 1986.
258. L. C. Zhao, P. R. Krishnaiah, and Z.-D.Bai. Remarks on certain criteria for detection of number of signals. *IEEE Trans. Acoust., Speech, Signal Processing* ASSP-35:129–132; 1987.
259. X. Zhuang, R. B. Harolick and Y. Zhao. Maximum entropy image reconstruction. *IEEE Trans. Acoust., Speech and Signal Processing* ASSP-39:1478–1480; 1991.

APPENDIX

LINEAR PREDICTION ESTIMATION

This Appendix gives an exposition of linear prediction estimation methods used in this review. For more rigorous and general treatments of linear prediction and its other applications, one is referred to (117,151,152).

Definitions

A data sequence $\{d_n\}$ is called p^{th}-order linear predictable if any element in it can be expressed as a *linear combination* of its previous (or next) p elements with a fixed set of weightings. In the forward prediction form,

$$d_n = -\sum_{l=1}^{p} h_l d_{n-l} \tag{A.1}$$

where the weightings $\{h_l; l = 1, \cdots, p\}$ are called the linear predictor coefficients and p is the predictor order. In the statistical literature, this data sequence is called p^{th}-order autoregressive (AR). One may notice that these terminologies have been used interchangeably throughout this review. A linear predictable sequence $\{d_n\}$ has many important properties; among them the following two have been exploited by various reconstruction algorithms discussed in this review.

- d_n can be generated from an all-pole system with the transfer function specified in the z-transform by

$$H(z) = \frac{1}{1 + \sum_{l=1}^{p} h_l z^{-l}}, \tag{A.2}$$

  or more precisely, d_n is the impulse response of $H(z)$.
- d_n can be expressed as a summation of exponentials[17] as

[17]This is not true if d_n is generated from an all-pole model with repeated poles; but this is not the case with the applications discussed in this review.

$$d_n = \sum_{l=1}^{p} a_l e^{-(\alpha_l + i\beta_l)n} \tag{A.3}$$

and $z_l = e^{-(\alpha_l + i\beta_l)}$, $l = 1, \cdots, p$, are the system poles of $H(z)$. The linear predictor is unstable if $H(z)$ has any poles outside the unit-circle in the complex z-plane, i.e., $\alpha_l < 0$, $\exists l$.

Applications of LP to image reconstruction in this review involve finding the coefficients h_l or the system poles z_l. Usually, z_l can be obtained from the LP characteristic polynomial, $\mathcal{P}(z) = 1 + \Sigma_{l=1}^{p} h_l z^{-l}$, after $\{h_l\}$ are known. Methods are also available to directly extract the z_l from the data, avoiding the polynomial rooting step. Several popular methods for these purposes are reviewed below.

Autocorrelation/Covariance Matrix Methods

In practice, the data sequence $\{d_n\}$ is not exactly linear predictable and the LP coefficients h_l are determined by minimizing the total squared linear prediction error, i.e.,

$$\min_{\{h_l\}} \sum_n e_n^2 = \min_{\{h_l\}} \sum_n \left(d_n + \sum_{l=1}^{p} h_l d_{n-l} \right)^2, \tag{A.4}$$

which leads to the following normal equations

$$\sum_{l=1}^{p} h_l \sum_n d_{n-m}^* d_{n-l} = -\sum_n d_{n-m}^* d_n, \qquad 1 \leq m \leq p. \tag{A.5}$$

If the length of the data sequence $\{d_n\}$ is infinite, Eq. (A.5) reduces to the well-known Yule-Walker equation

$$\sum_{l=1}^{p} h_l R(m - l) = -R(m), \qquad 1 \leq m \leq p, \tag{A.6}$$

where

$$R(m) = \sum_n d_n^* d_{n+m} \tag{A.7}$$

is the autocorrelation function of sequence $\{d_n\}$. In practice, the data length, N, is finite; in this case, the following biased estimate of the autocorrelation function is used

$$\hat{R}(m) = \sum_{n=0}^{N-1-m} (w_n d_n)^* (w_{n+m} d_{n+m}) \tag{A.8}$$

where w_n is any of the window functions discussed in (89). This method is called the *autocorrelation matrix* method in the signal processing literature; the autocorrelation matrix $\boldsymbol{R} = [R(M-l)]_{p\times p}$ is a Hermitian Toeplitz matrix and Eq. (A.6) can be solved very efficiently using the well-known Levinson-Durbin algorithm (27,122).

If the summation in Eq. (A.5) is taken over the known data interval ($0 \leq n \leq N-1$), one has

$$\sum_{l=1}^{p} h_l C_{ml} = -C_{m0}, \qquad 1 \leq m \leq p, \tag{A.9}$$

where

$$C_{ml} = \sum_{n=0}^{N-1} d_{n-l} d_{n-m}^* \tag{A.10}$$

is the covariance of the data sequence in the given interval and thus the method is referred to as the *covariance matrix* method. Unlike the autocorrelation matrix $\boldsymbol{R}$, the covariance matrix $\boldsymbol{C} = [C_{lm}]_{p\times p}$ is not Toeplitz and the fast Levinson-Durbin algorithm is not applicable, but since $\boldsymbol{C}$ is Hermitian and semi-definite the Cholesky method can be used to gain computational efficiency over the Gaussian elimination method. The main properties of these two methods are as follows.

- Autocorrelation matrix method:
  - — efficient (Levinson-Durbin algorithm);
  - — the resultant predictor is stable;
  - — resolution may be degraded by data windowing.
- Covariance matrix method:
  - — better resolution;
  - — less efficient (Cholesky algorithm);
  - — the resultant predictor may not be stable.

Of these two methods, the autocorrelation matrix method is more useful for LP extrapolation and ARMA modeling[18] for reconstruction stability. The covariance matrix method is more suitable for LPA modeling for higher resolution, since instability of the predictor may not result in reconstruction artifacts in this model. Note also that the covariance matrix method is identical to the common least squares method if the total prediction error e_n is minimized over the range $p+1 \leq n \leq N-1$.

[18]Smith et al. actually use Barrowdale's implementation of the covariance matrix method for estimating the AR parameters and handle the instability problem afterward.

SVD-Based Methods

SVD (Singular Value Decomposition) was introduced to LP parameter estimation for improved parameter accuracy with low SNR data by Kumaresan and Tufts (124,125,236). Their seminal work stimulated a tremendous amount of research in this area in the 80's, as evidenced by the number of papers published. This section aims to do no more than review some fundamental SVD-based techniques closely related to this review.

The Kumaresan-Tufts method. The Kumaresan-Tufts (KT) method [known as the LPSVD method in MR literature (11–13)] was originally developed for retrieval of the system poles $\{z_l, l = 1, \cdots, p\}$. This method obtains better parameter estimates than the autocorrelation matrix method or the covariance matrix method by the following two important steps.

- *Use of higher-order linear predictors:* For a p^{th}-order linear predictable sequence corrupted by noise, Kumaresan et al. suggested that a higher order ($\hat{p} \sim N/2 > p$) linear predictor be fitted to the data such that

$$\begin{bmatrix} d_0 & d_1 & \cdots & d_{\hat{p}-1} \\ d_1 & d_2 & \cdots & d_{\hat{p}} \\ \cdot & & & \cdot \\ \cdot & \cdots & \cdots & \cdot \\ \cdot & & & \cdot \\ d_{N-\hat{p}-1} & d_{N-\hat{p}} & \cdots & d_{N-2} \end{bmatrix} \begin{bmatrix} h_{\hat{p}} \\ h_{\hat{p}-1} \\ \cdot \\ \cdot \\ \cdot \\ h_1 \end{bmatrix} = - \begin{bmatrix} d_{\hat{p}} \\ d_{\hat{p}+1} \\ \cdot \\ \cdot \\ \cdot \\ d_{N-1} \end{bmatrix}, \quad \text{(A.11)}$$

or

$$\boldsymbol{D}\boldsymbol{h} = -\boldsymbol{d} \quad \text{(A.12)}$$

where $\boldsymbol{D}$ is a $(N - \hat{p}) \times \hat{p}$ linear prediction matrix. In the noiseless case, $\boldsymbol{D}$ is of rank p and the extra coefficients $h_{p+1}, \cdots, h_{\hat{p}}$ are zero. In the presence of noise, these coefficients (and the additional system poles thus introduced) can help to absorb noise.

- *Numerical filtering through the truncated SVD:* The other, more important, aspect of the Kumaresan-Tufts method is the replacement of the linear prediction matrix $\boldsymbol{D}$ by a new one with the desired rank p so that noise in $\boldsymbol{D}$ can be partly removed. This is accomplished through the use of the truncated SVD since it is the most effective numerical method for finding the best lower-rank approximation to a matrix. Represent $\boldsymbol{D}$ in the SVD notation,

$$\boldsymbol{D} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H = \sum_{l=1}^{\hat{p}} \lambda_l \boldsymbol{u}_l \boldsymbol{v}_l^H, \quad \text{(A.13)}$$

where $\boldsymbol{u}_l$ and $\boldsymbol{v}_l$ are the left and right singular vectors and $\Sigma = \text{diag}[\lambda_1, \cdots, \lambda_{\hat{p}}]$ with $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_{\hat{p}}$. It can be justified that the vector space spanned by the p most significant singular vectors ($\boldsymbol{u}_l$ and $\boldsymbol{v}_l$, $l = 1, \cdots, p$) corresponds to the signal

subspace[19] and the vector space spanned by the $\hat{p} - p$ least significant singular vectors corresponds to the noise subspace. The noise components in the noise subspace can then be removed by setting $\lambda_{p+1}, \cdots, \lambda_{\hat{p}}$ to zero, resulting in a new coefficient matrix[20] $\boldsymbol{D}_p$ with rank p

$$\boldsymbol{D}_p = \boldsymbol{U}_p \boldsymbol{\Sigma}_p \boldsymbol{V}_p^H = \sum_{l=1}^{p} \lambda_l \boldsymbol{u}_l \boldsymbol{v}_l^H. \tag{A.14}$$

The minimum-norm LS solution for $\boldsymbol{h}$ is given by

$$\boldsymbol{h}_{\mathrm{KT}} = \sum_{l=1}^{p} \frac{1}{\lambda_l} (\boldsymbol{u}_l^H \boldsymbol{d}) \boldsymbol{v}_l. \tag{A.15}$$

Note that the KT linear predictor defined in Eq. (A.15) introduces $\hat{p}$ - p extraneous poles. These poles are (approximately) uniformly distributed inside and along the unit circle and often do not cause a serious instability problem. Nonetheless, the KT linear predictor is not guaranteed to be stable in general and removal of the unstable extraneous poles may be necessary to avoid reconstruction artifacts, depending on the application.

Total least squares (TLS) method. In the KT method, noise in the right hand side vector $\boldsymbol{d}$ is not subject to numerical filtering by the truncated SVD and will perturb the accuracy of the estimated predictor coefficients. Moving the right-hand-side data vector into the coefficient matrix to create the following homogeneous equation:

$$[\boldsymbol{D} \vdots \boldsymbol{d}] \begin{bmatrix} \boldsymbol{h} \\ \cdots \\ 1 \end{bmatrix} = 0, \tag{A.16}$$

noise in $\boldsymbol{D}$ as well as in $\boldsymbol{d}$ will be filtered by applying the truncated-SVD to this new augmented coefficient matrix $\tilde{\boldsymbol{D}}$ and, thus, the name *total least squares* method[21] (183,232). Using the notation of Eq. (A.14), the truncated coefficient matrix is

$$\tilde{\boldsymbol{D}}_p = \boldsymbol{U}_p \boldsymbol{\Sigma}_p \boldsymbol{V}_p^H = \sum_{l=1}^{p} \lambda_l \boldsymbol{u}_l \boldsymbol{v}_l^H. \tag{A.17}$$

Eq. (A.16) then becomes

[19]This may not be true when the data SNR is below a threshold, in which case, truncating the SVD will lead to a loss of signal components.

[20]Note that the Hankel structure is not preserved in this matrix; some iterative methods for finding a lower-rank Hankel approximation have been discussed in the recent signal processing literature.

[21]This method is also known as the generalized Pisarenko method or noise eigenvector method, recognizing that $\tilde{\boldsymbol{D}}$ is in fact a $(N - \hat{p}) \times (\hat{p} + 1)$ linear prediction matrix and $[\boldsymbol{h}, 1]^t$ is one of its noise eigenvectors.

$$V_p^H \begin{bmatrix} \boldsymbol{h} \\ \cdots \\ 1 \end{bmatrix} = 0, \tag{A.18}$$

or

$$\boldsymbol{C}\boldsymbol{h} = -\boldsymbol{c}, \tag{A.19}$$

where

$$C = \begin{bmatrix} v_{11}^* & v_{21}^* & \cdots & v_{p1}^* \\ v_{12}^* & v_{22}^* & \cdots & v_{p2}^* \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & & \cdot \\ v_{1p}^* & v_{2p}^* & \cdots & v_{pp}^* \end{bmatrix}, \quad \text{and} \quad \boldsymbol{c} = \begin{bmatrix} v_{(p+1)1}^* \\ v_{(p+1)2}^* \\ \cdot \\ \cdot \\ \cdot \\ v_{(p+1)p}^* \end{bmatrix}.$$

The minimum-norm LS solution for $\boldsymbol{h}$ is

$$\boldsymbol{h}_{\mathrm{TLS}} = -\boldsymbol{C}^H(\boldsymbol{C}\boldsymbol{C}^H)^{-1}\boldsymbol{c}. \tag{A.20}$$

Making use of the identity

$$(\boldsymbol{C}\boldsymbol{C}^H)^{-1} = (\boldsymbol{I} - \boldsymbol{c}\boldsymbol{c}^H)^{-1} = \boldsymbol{I} + \frac{\boldsymbol{c}\boldsymbol{c}^H}{1 - \|\boldsymbol{c}\|^2} \tag{A.21}$$

one has

$$\boldsymbol{h}_{\mathrm{TLS}} = -\boldsymbol{C}^H \left(\boldsymbol{I} + \frac{\boldsymbol{c}\boldsymbol{c}^H}{1 - \|\boldsymbol{c}\|^2} \right) \boldsymbol{c}, \tag{A.22}$$

or

$$\boldsymbol{h}_{\mathrm{TLS}} = -\frac{\boldsymbol{C}^H\boldsymbol{c}}{1 - \|\boldsymbol{c}\|^2}. \tag{A.23}$$

An equivalent expression can also be derived in terms of the noise singular vectors as

$$\boldsymbol{h}_{\mathrm{TLS}} = \frac{\tilde{\boldsymbol{C}}^H\tilde{\boldsymbol{c}}}{\|\boldsymbol{c}\|^2} \tag{A.24}$$

where

$$\tilde{\boldsymbol{C}} = \begin{bmatrix} v^*_{1(p+1)} & v^*_{2(p+1)} & \cdots & v^*_{\hat{p}(p+1)} \\ v^*_{1(p+2)} & v^*_{2(p+2)} & \cdots & v^*_{\hat{p}(p+2)} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & & \cdot \\ v^*_{1(\hat{p}+1)} & v^*_{2(\hat{p}+1)} & \cdots & v^*_{\hat{p}(\hat{p}+1)} \end{bmatrix}, \quad \text{and} \quad \tilde{\boldsymbol{c}} = \begin{bmatrix} v^*_{(\hat{p}+1)(p+1)} \\ v^*_{(\hat{p}+1)(p+2)} \\ \cdot \\ \cdot \\ \cdot \\ v^*_{(\hat{p}+1)(\hat{p}+1)} \end{bmatrix}$$

From expressions (A.23) or (A.24), one can see that the TLS method has a computational advantage over the KT method since it requires only the right singular vectors and, thus, eliminates the need for computing the full SVD.

State-space method. The state-space formulation of the linear prediction problem was initially developed by Kung, Arun, and Rao (126). A more detailed and rigorous discussion of this method can be found in a recent review article (185). This method has been used for NMR spectroscopic data processing, where it is called the HSVD method (14). Following the derivation by Yan and Gore (249), one can treat the LP parameter estimation problem in this formulation as finding a matrix $\boldsymbol{Z}$ satisfying

$$\boldsymbol{D}_b\boldsymbol{Z} = \boldsymbol{D}_t, \tag{A.25}$$

where $\boldsymbol{D}_b$ and $\boldsymbol{D}_t$ are the matrices obtained by removing the bottom and top rows of the following linear prediction matrix $\boldsymbol{D}$,

$$\boldsymbol{D} = \begin{bmatrix} d_0 & d_1 & \cdots & d_{\hat{p}-1} \\ d_1 & d_2 & \cdots & d_{\hat{p}} \\ \cdot & & & \cdot \\ \cdot & \cdots & \cdots & \cdot \\ \cdot & & & \cdot \\ d_{N-\hat{p}} & d_{N-\hat{p}+1} & \cdots & d_{N-1} \end{bmatrix}. \tag{A.26}$$

The solution of $\boldsymbol{Z}$ is given by (249)

$$\boldsymbol{Z} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -h_{\hat{p}} \\ 1 & 0 & 0 & \cdots & 0 & -h_{\hat{p}-1} \\ 0 & 1 & 0 & \cdots & 0 & -h_{\hat{p}-2} \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & -h_1 \end{bmatrix}, \tag{A.27}$$

where the h_l are defined through Eq. (A.11). Since the characteristic polynomial of $\boldsymbol{Z}$ is

$$\det(z\boldsymbol{I} - \boldsymbol{Z}) = \det \begin{bmatrix} z & 0 & 0 & \cdots & 0 & h_{\hat{p}} \\ -1 & z & 0 & \cdots & 0 & h_{\hat{p}-1} \\ 0 & -1 & z & \cdots & 0 & h_{\hat{p}-2} \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -1 & z + h_1 \end{bmatrix} = z^{\hat{p}} + h_1 z^{\hat{p}-1} + \cdots + h_{\hat{p}}, \tag{A.28}$$

the eigenvalues of $\boldsymbol{Z}$ provide exactly the desired roots of the LP characteristic polynomial, but without implicitly solving for the polynomial coefficients h_l and the rooting step as it is done in the KT or TLS methods. One further important result is that $\boldsymbol{D}_b$ and $\boldsymbol{D}_t$ in Eq. (A.25) can be replaced by the truncated left or right singular matrix of $\boldsymbol{D}$ without affecting the above property of $\boldsymbol{Z}$ (249). For example, if the left truncated singular matrix $\boldsymbol{U}_p$ is used, the solution of $\boldsymbol{Z}$ from the following equation

$$\boldsymbol{U}_b\boldsymbol{Z} = \boldsymbol{U}_t \tag{A.29}$$

is given by (14,134)

$$\boldsymbol{Z} = \left(\boldsymbol{I} + \frac{\tilde{\boldsymbol{u}}^H_{N-\hat{p}+1}\hat{\boldsymbol{u}}_{N-\hat{p}+1}}{1 - \|\hat{\boldsymbol{u}}_{N-\hat{p}+1}\|^2}\right)\boldsymbol{U}_b^H\boldsymbol{U}_t, \tag{A.30}$$

where $\hat{\boldsymbol{u}}_{N-\hat{p}+1}$ is the $(N - \hat{p} + 1)^{th}$ row vector of $\boldsymbol{U}_p$ formed from the p principal left singular vectors of $\boldsymbol{D}$. A similar expression exists for the right singular vectors. The $\boldsymbol{Z}$ matrix defined above has p eigenvalues, and they are exactly equal to the p signal roots of the KT linear predictor in the noiseless case. Therefore, one appealing feature of this state-space approach is that it directly gives the system poles (or signal roots) and thus no further measures are necessary to deal with the extraneous poles present in both the KT and TLS methods. Also, similar to the TLS method, this method only requires the left (or right) singular vectors and thus eliminates the need to compute the full SVD of the linear prediction matrix $\boldsymbol{D}$.

Combined forward-backward linear prediction. The LP techniques described above are all formulated in terms of the forward linear prediction. One can derive similar expressions for backward linear prediction. A more useful formulation results when the forward and backward linear prediction (FBLP) are combined to give (124)

$$\begin{bmatrix} d_0 & d_1 & \cdots & d_{\hat{p}-1} \\ d_1 & d_2 & \cdots & d_{\hat{p}} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ d_{N-\hat{p}-1} & d_{N-\hat{p}} & & d_{N-2} \\ \hline d^*_{\hat{p}} & d^*_{\hat{p}-1} & \cdots & d^*_1 \\ d^*_{\hat{p}+1} & d^*_{\hat{p}} & \cdots & d^*_2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ d^*_{N-1} & d^*_{N-2} & \cdots & d^*_{N-\hat{p}} \end{bmatrix} \begin{bmatrix} h_p \\ h_{p-1} \\ \cdot \\ \cdot \\ \cdot \\ h_1 \end{bmatrix} = - \begin{bmatrix} d_p \\ d_{p+1} \\ \cdot \\ \cdot \\ \cdot \\ d_{N-1} \\ \hline d^*_0 \\ d^*_1 \\ \cdot \\ \cdot \\ \cdot \\ d^*_{N-\hat{p}-1} \end{bmatrix}. \tag{A.31}$$

The most important property of this formulation is its ability to constrain the system poles onto the unit circle. It can lead to a significant improvement in the parameter estimates if the data are known to consist of a set of undamped sinusoids, such as is the case with LPA modeling. Otherwise, it may cause model instability, hence, reconstruction artifacts (notably, noise spikes), as it does with LP extrapolation and ARMA modeling. Note that if the data sequence $\{d_n\}$ is symmetric, Eq. (A.11) is equivalent to a FBLP formulation if both wings of the data are used, and, therefore, the formulation in Eq. (A.31) is not

necessary. For this reason, in ARMA modeling and LP extrapolation, each wing of the data needs to be modeled separately to avoid instability problems.[22]

NOMENCLATURE

$\rho(x)$:	True image function.
$\hat{\rho}(x)$:	Estimated image function.
$\rho_{ARMA}(x)$:	ARMA model function.
$\rho_{GS}(x)$:	GS model function.
$\rho_{LPA}(x)$:	LPA model function.
$\rho_{DFT}(x)$:	DFT reconstruction.
$\rho_{WFT}(x)$:	Windowed FT reconstruction.
$\rho_{slim}(\boldsymbol{x},f)$:	SLIM spectral function.
$\rho_{gslim}(\boldsymbol{x},f)$:	GSLIM spectral function.
$\rho_m^{slim}(f)$:	SLIM compartmental spectrum.
$\rho_m^{gslim}(f)$:	GSLIM compartmental spectrum.
$\rho_m^{ft}(f)$:	FT compartmental spectrum.
$s(k)$:	Noise-free Fourier signal.
$\hat{s}(k)$:	Noise-corrupted version of $s(k)$.
$\bar{s}(k)$:	Processed version of $s(k)$, for example, $\bar{s}(k) = i2\pi k s(k)$.
$s(\boldsymbol{k},t)$:	Phase-encoded spectroscopic MR signal.
$\hat{s}(\boldsymbol{k},t)$:	Noise-corrupted version of $s(\boldsymbol{k},t)$.
$\bar{s}(\boldsymbol{k},f)$:	Temporal Fourier transform of $s(\boldsymbol{k},t)$.
$\varphi(x)$:	True phase image; $\varphi(x) = \arg\{\rho(x)\}$.
$\hat{\varphi}(x)$:	An estimate of $\hat{\varphi}(x)$.
$\varphi_{DFT}(x)$:	DFT reconstruction of $\varphi(x)$.
$\mathcal{N}_{data}$:	Integer set of indices for sampling of the Fourier signal.
N_{data}:	Total number of data points collected; $\|\mathcal{N}_{data}\| = N_{data}$.
$\mathcal{N}_{imag}$:	Integer set of indices for sampling of the image function.
N_{imag}:	Total number of image pixels; $\|\mathcal{N}_{imag}\| = N_{imag}$.
$\boldsymbol{G}$:	SLIM geometry matrix.
$\kappa(\boldsymbol{G})$:	Condition number of $\boldsymbol{G}$.
$\boldsymbol{G}^+ = [g_{mn}^+]$:	Pseudo-inverse of $\boldsymbol{G}$.
$\boldsymbol{E}$:	Extrapolation matrix.
$\boldsymbol{I}$:	Identity matrix.
$\mathcal{B} = (-L_x/2, L_x/2)$:	Spatial support bound for $\rho(x)$.
H_S:	Shannon entropy measure.
H_B:	Burg entropy measure.
Δk:	Spatial-frequency sampling interval.
Δx, Δx_d:	Image digital pixel size.
δx, δx_{opt}:	Effective width of point spread function.
Δx_f:	Fourier pixel size, $\Delta x_f = \delta x_{opt} = 1/(N_{data}\Delta k)$.
$h(x)$:	Point spread function.
$\mathcal{P}_m(x)$:	Contributing function for compartmental localization.
$\mathcal{P}(z)$:	LP characteristic polynomial.
$w(x)$ or $w(k)$:	Window function defined based on the contexts.

[22] In ARMA modeling, Smith et al. uses an alternative method to overcome this problem; see the section on the LPA model for more details.

$\varepsilon(k)$, $\epsilon(n)$, $e(n)$:	Noise terms.
$\mathcal{C}(\boldsymbol{\theta},x)$:	Constraint function in the GS model.
λ:	Eigenvalue (or singular) value of a matrix.
μ:	Regularization parameter or Lagrange multiplier.
γ:	Gyromagnetic ratio.
σ:	Noise standard deviation.
Δ_p:	Amplitude thresholding parameter of the sigma-filter.
$\mathcal{A}_m$:	SLIM noise amplification factor.
$\mathcal{D}_m$:	SLIM compartmental geometry.
$\mathcal{L}_{mn}$:	SLIM inter-compartmental spectral leakage.
$E\{\bullet\}$:	Statistical expectation value.
$\mathcal{F}\{\bullet\}$, $\mathcal{F}^{-1}\{\bullet\}$:	Forward or inverse FFT.
arg{•}:	Phase from a complex number.
Var{•}:	Variance of a random variable.
$\lvert\bullet\rvert$:	Absolute value of a real or complex number.
$\lVert\bullet\rVert$:	Absolute value of a vector or matrix.
det{•}:	Determinant of a matrix.
Superscript *:	Complex conjugate of a complex variable, matrix or vector.
Superscript *H*:	Hermitian transpose of a matrix or vector.
Superscript *t*:	Transpose of a matrix or vector.
$\forall$:	Meaning "for all."
$\exists$:	Meaning "there exists."
{•}:	Meaning "the set (or sequence) of."
$\lvert$:	Meaning "under the conditions of or such that."
$\in$:	Meaning "an element of."
$\cap$:	Set intersection operator.
$\star$:	Convolution operator.

GLOSSARY

Aliasing: An image artifact arising from undersampling of the Fourier space, which manifests itself as folding-over (ghosting) of some parts of the object in the Fourier reconstruction.

Bandlimited function: A function is said to be bandlimited if its Fourier transform has non-zero values only inside a finite frequency band.

Bandwidth: Frequency interval over which the values of the Fourier transform of a bandlimited function are non-zero.

Chemical shift imaging: Often used to refer to a special class of spectroscopic imaging techniques that use phase-encoding or frequency-encoding to encode the spatial origins of spectroscopic signals.

Condition number: A measure of the singularity of a matrix. In the spectral norm, it is defined as the ratio of the largest eigenvalue over the smallest one.

Contractive operator: An operator φ is called contractive if it reduces the norm (magnitude), i.e., $\|\varphi x\| < \|x\|$.

Convex set: Set with the property that points on the line joining any two of its members are inside this set.

Convolution: Mathematical operation between two functions, denoted in this review by "$\star$" and defined through the relation $(f \star g)(t) = \int_{-\infty}^{\infty} f(\mathrm{r})g(r - t)dt$.

Convolution kernal: In $f \star g$, g is called the convolution kernel for f or f is the convolution kernel for g depending on the context. For example, if $f \star g$ is taken as the output of a linear system within input f, then g is the kernel function, which is also called *transfer function* in system theory and *point spread function* in imaging, because $f \star g = g$ if $f = \delta$, a point source.

Ergodic process: A stochastic process in which ensemble averages are equivalent to time sample averages.

Expansive operator: An operator φ is called expansive if it amplifies the norm (magnitude), i.e., $\|\varphi x\| > \|x\|$.

Extrapolation: The mathematical process of extending the definition domain of a function.

FID (Free Induction Decay): Transient NMR signal observed following radio-frequency excitation pulses.

Finite sampling: Process of collecting a finite number of discrete samples or measurements from a continuous function or a physical process.

Frequency-encoding: Process of establishing a correspondence between the spatial location and the precessional frequency of an MR signal.

Gibbs' artifact: An artifact associated with the approximation of a discontinuous function by a truncated Fourier series. This artifact is characterized by spurious ringing near the area of discontinuities.

Ill-conditioned: A matrix (or more generally, any operator) is called ill-conditioned if it has a large condition number κ; the two important extreme cases are $\kappa = 1$ and $\kappa = \infty$, the former corresponding to an orthogonal (best-conditioned) operator and the latter corresponding to a singular (worst-conditioned) operator. An ill-conditioned operator often means that the solution of the corresponding operator equation, for example, a linear equation, is sensitive to perturbations (noise). Procedures to reduce this sensitivity are often called *regularization*.

***k*-space:** The spatial frequency domain where the magnetic resonance imaging signals are collected.

MRI (Magnetic Resonance Imaging): Methods whereby NMR signals are spatially encoded using magnetic field gradients so that the spatial distribution of NMR properties, such as spin density and relaxation times, of a sample can be reconstructed and presented in an image format.

MRS (Magnetic Resonance Spectroscopy): A method to obtain NMR spectra displaying chemically-shifted resonance peaks that provide information on the chemical species present in the system and their relative concentrations.

MR spectroscopic imaging: MR imaging techniques to obtain spatially-resolved MR spectra.

Nyquist frequency: The minimum sampling frequency to guarantee an exact reconstruction of a continuous function from an infinite set of discretely sampled values. Aliasing will result if the sampling frequency is below the Nyquist frequency.

Periodic extension: Periodic extension of a $[-L_x/2, L_x/2]$-support-limited function $\rho(x)$ with a period $\tilde{L}_x = 1/\Delta k \geq L_x$ is defined as

$$\begin{cases} \tilde{\rho}(x) = \rho(x), & x \in [-L_x/2, L_x/2] \\ \tilde{\rho}(x) = 0, & -\tilde{L}_x/2 \leq x < -L_x/2 \text{ or } L_x/2 < x \leq \tilde{L}_x/2 \\ \tilde{\rho}(x) = \tilde{\rho}(x + n\tilde{L}_x), & \forall n. \end{cases}$$

Phase Artifact: Often used in connection with the image artifact associated with a partial Fourier reconstruction resulting from phase inconsistency between the extrapolated data and the original data.

Phase-encoding: Process of establishing a correspondence between the spatial location and the phase of an MR signal.

Point spread function: see the definition for *convolution kernel.*

Projection operator: Any operator φ having the property that $\varphi^2 = \varphi$. A projection operator is said to be orthogonal if $\|\varphi x - x\|$ is a minimum.

Pseudo-inverse: A generalized definition of matrix (or operator) inverse for singular or non-square matrices.

QR decomposition: Decomposition of a matrix into the product $\boldsymbol{QR}$ of an orthonormal matrix $\boldsymbol{Q}$ and an upper triangular matrix $\boldsymbol{R}$.

Regularization: See *Ill-conditioned.*

Spectral localization: Obtaining MR spectra from selected regions.

Superresolution: In Fourier imaging, super-resolution refers to the ability to resolve image features with separations smaller than the Fourier pixel size.

Support-limited function: A function is said to be support-limited if it has non-zero values only inside a finite spatial region, in contrast to a band-limited function whose Fourier transform function has non-zero values inside a finite frequency band. Note that a function cannot be both support-limited and bandlimited at the same time.

Support: Spatial interval over which a support-limited function has non-zero values, similar to bandwidth for bandlimited functions.

SVD (Singular Value Decomposition): Decomposition of a matrix to the form $\boldsymbol{U\Sigma V}^{\mathrm{H}}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthonormal matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix.

Toeplitz matrix: Matrix for which all the elements along any diagonal are identical.

Trunctaion artifact: A reconstruction artifact due to the loss of high spatial frequency data. With the Fourier reconstruction method, it manifests itself as image blurring and spurious ringing near sharp edges. The data truncation artifact may behave differently in different image reconstruction methods.

Variable projection optimization: An optimization technique associated with the solution of certain nonlinear least squares problems where the model has a nonlinear as well as a linear dependence on its parameters. It improves computational throughput over standard nonlinear optimization techniques by decoupling the calculation of the nonlinear parameters from the linear ones.

PHASE IMAGES

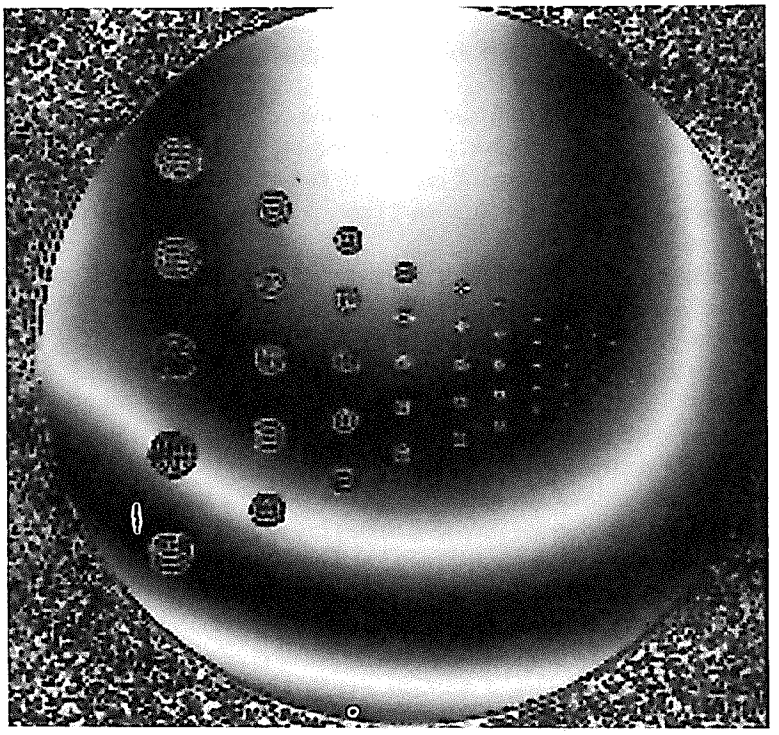

Figure A1. Phase image of the phantom data set.

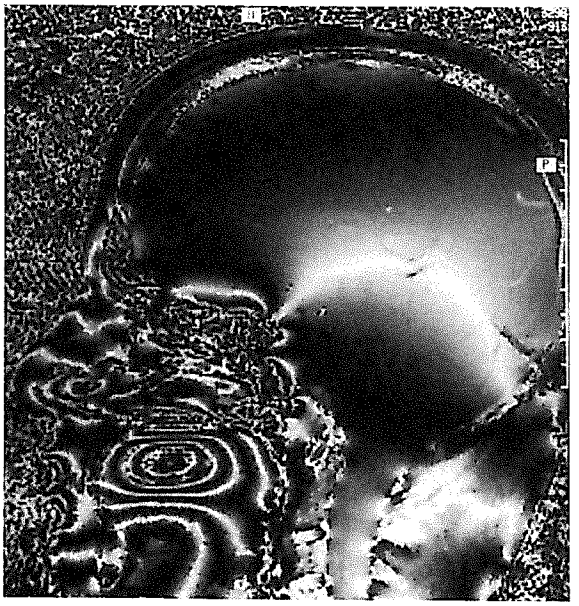

Figure A2. Phase image of the sagittal head data set.

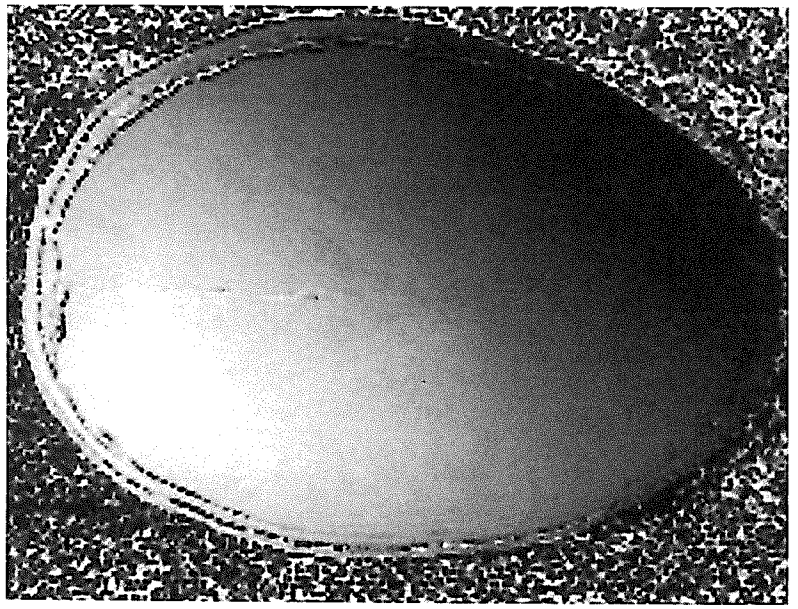

Figure A3. Phase image of the transverse head data set.

Figure A4. Phase image of the sagittal lumbar spine data set.