## **Chapter 6**

# **Reconstruction from Fourier Samples** (Gridding and alternatives)

ch, fou

## Contents

6.1	Introd	uction (s,four,intro)
	6.1.1	Problem statement (s,four,state)
	6.1.2	Solution strategies
	6.1.3	Space-limited objects
	6.1.4	Operator formulation (s,four,op)
		6.1.4.1 Adjoint operator
		6.1.4.2 Frame operator
		6.1.4.3 Crosstalk matrix
		6.1.4.4 Weighted crosstalk matrix
6.2	Finite-	series (discrete-discrete) methods (s,four,series)
	6.2.1	Finite-series object model
	6.2.2	Weighted least-squares (WLS) solution
	6.2.3	Penalized weighted least-squares (PWLS) method
	6.2.4	Choosing the regularization parameter
	6.2.5	Equally-spaced basis functions
	6.2.6	Choice of basis functions
	6.2.7	Uniform frequency samples case
	6.2.8	Complex exponential basis
	6.2.9	Conjugate-gradient algorithm for practical PWLS estimation (s,four,series,pcg)
	6.2.10	Toeplitz embedding
	6.2.11	Effect of number of basis functions and pixel size (s,four,series,pixel)
	6.2.12	General linear reconstructor (s,four,series,lin)
	6.2.13	Maximum entropy formulations (s,four,maxent)
	6.2.14	Sparse reconstruction (s,four,sparse)
6.3	Discre	te-data, continuous-object methods (s,four,dc)
	6.3.1	Minimum-norm least-squares methods (s,four,minnorm)
	6.3.2	Penalized least-squares for discrete data, continuous object model (s,four,pls)
	6.3.3	Relation between MNLS and QPWLS (s,four,dc,mnls)
6.4	Contir	uous-continuous methods (s,four,cc)
	6.4.1	Conjugate phase method (s,four,conj)
	6.4.2	Sampling density compensation (s,four,dens)
		6.4.2.1 Noise considerations (s,four,dens,noise)
		6.4.2.2 Heuristic methods (s,four,dens,heur)
		6.4.2.2.1 Jacobian determinant

		6.4.2.2.2 Voronoi cell volume	6.18
		6.4.2.2.3 Cell counting	6.18
		6.4.2.2.4 Sinc overlap density	6.18
		6.4.2.2.5 Jackson's area density	6.19
		6.4.2.3 Optimality criteria: image domain (s,four,dens,im)	6.20
		6.4.2.3.1 Object-dependent image-domain DCF	6.20
		6.4.2.3.2 Min-max image-domain DCF	6.20
		6.4.2.3.3 Weighted Frobenius norm criterion in data domain	6.21
		6.4.2.3.4 Weighted Frobenius norm criterion in image domain	6.21
		6.4.2.3.5 Pseudo-inverse criteria (s,four,dens,pinv)	6.22
		6.4.2.4 Optimality criteria: PSF (s,four,dens,psf)	6.23
		6.4.2.4.1 Pipe and Menon's iteration	6.24
		6.4.2.4.2 An EM iteration	6.25
		6.4.2.4.3 Qian et al.'s iteration	6.25
		6.4.2.4.4 Samsonov et al.'s iteration	6.25
		6.4.2.4.5 Alternative PSF-based methods	6.25
		6.4.2.5 Optimality criteria: frequency domain (s,four,dens,freq)	6.26
		6.4.2.5.1 Weighted Frobenius norm in data domain	6.26
		6.4.2.6 Density compensation summary (s,four,dens,summ)	6.27
	6.4.3	Frequency-domain interpolation methods (s,four,freq)	6.28
		6.4.3.1 Block spectral interpolation methods (s,four,freq,block)	6.28
		6.4.3.2 The "block uniform resampling" (BURS) method	6.28
		6.4.3.3 Min-max spectral interpolation (s,four,freq,minmax)	6.29
		6.4.3.4 Iterative spectrum estimation (s,four,freq,iter)	6.30
		6.4.3.5 Gridding methods (s,four,grid)	6.30
		6.4.3.6 Practical gridding implementation (s,four,freq,impl)	6.32
6.5	Summ	ary (s,four,summ)	6.32
6.6	Appen	dix: NUFFT calculations (s,four,nufft)	6.35
	6.6.1	Basic ideal interpolator	6.35
	6.6.2	Generalized ideal interpolator	6.36
	6.6.3	Practical interpolators	6.36
6.7	Appen	dix: NUFFT-based gridding (s,four,nufft,grid)	6.38
6.8	Appen	dix: Toeplitz matrix-vector multiplication (s,four,toep)	6.42
6.9	Proble	ms (s,four,prob)	6.43
6.10	Bibliog	graphy	6.45

## 6.1 Introduction (s,four,intro)

There are a wide variety of inverse problems in which one can consider the available data to be samples of the Fourier transform of an object f of interest, and the goal is to reconstruct that object from those (noisy) samples. Examples include radio astronomy [1–5], radar [6–10], radiometers [11], ultrasound [12–14], magnetic resonance imaging [15–18], and tomographic image reconstruction (for 2D parallel-beam geometry) based on the Fourier slice theorem as described in §3.4. This chapter describes solution methods for such inverse problems.

Rarely is the sampled Fourier transform an exact model; for example, in MRI this formulation ignores many physical effects such as field inhomogeneity and spin relaxation. (See Chapter 7.) Nevertheless, it can be a useful starting point. This problem also allows us to introduce general principles for solving inverse problems without the complications of accurate physical models.

A related problem of interest in fields such as astronomy is that of performing spectral analysis of unevenly sampled data, *e.g.*, [19, 20]. Some methods of this chapter may be suitable for such applications.

Throughout this chapter we focus on methods that are applicable to general sampling patterns. There are also techniques for specific sampling patterns, such as spirals e.g., [21-23] and concentric rings [24]. In particular, the case of polar samples has been studied extensively due to its relevance to tomography, e.g., [25-34]. Several of those methods are based on polar to Cartesian interpolation schemes that are specific to the case of polar sampling. An interesting open question is whether the methods for general sampling have progressed sufficiently to make the pattern specific techniques unnecessary.

#### Ir, state 6.1.1 Problem statement (s, four, state)

Let  $f(\vec{x})$  denote the unknown object, *e.g.*, the patient's transverse magnetization in MRI, where  $\vec{x} \in \mathbb{R}^{\overline{d}}$ . Typically  $\overline{d} = 2$  or  $\overline{d} = 3$ , but the methods described apply to any dimension  $\overline{d} \in \mathbb{N}$ . Sometimes we write  $f^{\text{true}}$  rather than just f to denote the "true" unknown object. The available measurements are  $n_{\text{d}}$  samples of the Fourier transform of f at spatial frequencies  $\vec{\nu}_1, \ldots, \vec{\nu}_{n_{\text{d}}}$ , contaminated by measurement noise. Fig. 6.1.1 illustrates some of the frequency domain sampling patterns of interest in MRI. For simplicity we assume an **additive noise** model:

$$y_i = F(\vec{\nu}_i) + \varepsilon_i, \qquad i = 1, \dots, n_d, \tag{6.1.1}$$

where  $\varepsilon_i$  is zero-mean (complex) noise, and  $F(\vec{\nu})$  denotes the Fourier transform of f for  $\vec{\nu} \in \mathbb{R}^{\bar{d}}$ , defined as follows:

$$F(\vec{\nu}) = \int_{\mathbb{R}^{\bar{d}}} f(\vec{x}) e^{-i2\pi\vec{\nu}\cdot\vec{x}} \, \mathrm{d}\vec{x} \,. \tag{6.1.2}$$

Our goal is to estimate the true object function  $f^{\text{true}}$  from the measurement vector  $\boldsymbol{y} = (y_1, \dots, y_{n_d})$ .



Figure 6.1.1: Examples of 2D k-space sampling patterns  $\{\vec{\nu}_i\}$  of interest in MRI.

We assume throughout that the frequency sample locations  $\{\vec{\nu}_i\}$  are *distinct*. If the same spatial frequency were sampled more than once, then one could aggregate the corresponding measurements by averaging, thereby reducing the problem to the case of distinct frequencies.

The problem (6.1.1) involves *discrete* data but an unknown function f of *continuous* spatial variables. Such **discrete-continuous** problems seem to be hopelessly non-unique, because there are infinitely many functions  $\hat{f}$  that agree exactly with the measurements, *i.e.*, that satisfy the equalities  $y_i = \hat{F}(\vec{v}_i)$  for  $i = 1, ..., n_d$ . However, in the presence of noise, *none* of these possible "solutions" will match  $f^{\text{true}}$  exactly, and most of these "solutions" will be uselessly noisy. Thus, to "solve" this problem we must

- constrain f to lie within some set of useful functions,
- state a criterion for specifying which estimate f in that set is the "best" given the measurement vector y, and
- find a numerical algorithm for computing that f.

Interestingly, there are numerous publications related to **partial k-space** reconstruction methods, *e.g.*, [35, 36]. Strictly speaking, *any* finite set of Fourier samples is a "partial" set, because no finite set can uniquely determine f exactly, without further constraints, even in the absence of noise.

#### 6.1.2 Solution strategies

As summarized elegantly in [37] in the context of emission tomography, there are three general categories of solutions to inverse problems such as (6.1.1). Because the data y is finite dimensional, one option is to also discretize or parameterize the object f and estimate those parameters from y. This **discrete-discrete** approach is described in §6.2.

e,four,Fpu

Alternatively, one can tackle the **discrete-continuous** model (6.1.1) directly, as described in §6.3. Another option is to imagine hypothetical measurements that are indexed by a *continuous* set of variables, solve for f in using those hypothetical measurements, and then attempt to implement a practical discretization of that solution. (The FBP method for tomographic reconstruction is an example of this approach because it is based on the idealized Radon transform model for tomography.) This **continuous-continuous** formulation is particularly popular in MRI, and is described in §6.4.1 and §6.4.3.

#### 6.1.3 Space-limited objects

All of the methods described below assume that f is space-limited to some known subset  $\mathcal{D}$  of  $\mathbb{R}^{\bar{d}}$ . This assumption is reasonable physically (objects are space limited, except perhaps in astronomy), and it seems virtually essential from sampling theory because we are sampling in frequency space. In particular, in medical imaging  $\mathcal{D}$  corresponds to (at most) the field of view (FOV) of the scanning device, within which the object (*i.e.*, patient) must fit. For  $\bar{d} = 2$ , typically  $\mathcal{D} = [-\text{FOV}/2, \text{FOV}/2] \times [-\text{FOV}/2, \text{FOV}/2]$ , a square, or  $\mathcal{D} = \left\{ (x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} \le \text{FOV}/2 \right\}$ , a circle.

For space-limited functions, we can replace (6.1.2) with an integral over  $\mathcal{D}$ :

$$F(\vec{\nu}) = \int_{\mathcal{D}} f(\vec{\mathbf{x}}) e^{-i2\pi\vec{\nu}\cdot\vec{\mathbf{x}}} \, \mathrm{d}\vec{\mathbf{x}}$$

so

$$F(\vec{\nu}_i) = \langle f, \phi_i \rangle = \int_{\mathcal{D}} f(\vec{\mathbf{x}}) \, \phi_i^*(\vec{\mathbf{x}}) \, \mathrm{d}\vec{\mathbf{x}}, \tag{6.1.3}$$

where we define

$$\phi_i(\vec{\mathbf{x}}) \triangleq e^{i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}} \mathbb{I}_{\{\vec{\mathbf{x}} \in \mathcal{D}\}}, \qquad i = 1, \dots, n_d.$$
(6.1.4)

From (6.1.3), alternative expressions are

$$F(\vec{\nu}_i) = \int F(\vec{\nu}) S^*(\vec{\nu} - \vec{\nu}_i) \, \mathrm{d}\vec{\nu} = \int F(\vec{\nu}) S(\vec{\nu}_i - \vec{\nu}) \, \mathrm{d}\vec{\nu}$$
(6.1.5)

$$= (S * F)(\vec{\nu}_i),$$
 (6.1.6)

where \* denotes  $\bar{d}$ -dimensional convolution and  $S(\vec{\nu})$  denotes the  $\bar{d}$ -dimensional Fourier transform of the spatial support function  $s(\vec{x})$  that is defined by

$$s(\vec{\mathbf{x}}) \triangleq \mathbb{I}_{\{\vec{\mathbf{x}}\in\mathcal{D}\}} \xleftarrow{\mathrm{FT}} S(\vec{\nu}) = \int_{\mathcal{D}} e^{-i2\pi\vec{\nu}\cdot\vec{\mathbf{x}}} \,\mathrm{d}\vec{\mathbf{x}} \,. \tag{6.1.7}$$

Typically  $\mathcal{D}$  is square or circular, so  $S(\vec{\nu})$  is a sinc or jinc function. Because  $s(\vec{x})$  is real,  $S(\cdot)$  has conjugate symmetry:  $S(-\vec{\nu}) = S^*(\vec{\nu})$ .

#### s, four, op 6.1.4 Operator formulation (s, four, op)

It will simplify notation (and unify concepts) to express the model (6.1.1) using linear algebra. Defining the linear operator  $\mathcal{A} : \mathcal{L}_2(\mathcal{D}) \mapsto \mathbb{C}^{n_d}$  as follows:

$$[\mathcal{A}f]_i = \int f(\vec{\mathbf{x}}) \,\phi_i^*(\vec{\mathbf{x}}) \,\mathrm{d}\vec{\mathbf{x}},\tag{6.1.8}$$

rewrite (6.1.1) as

or equivalently as

$$\boldsymbol{y} = \boldsymbol{\mathcal{A}} f + \boldsymbol{\varepsilon}, \tag{6.1.9}$$

where  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_{n_d})$ . Loosely speaking, one can think of  $\mathcal{A}$  as a  $n_d \times \infty$  "matrix," conveying that this reconstruction problem is severely under-determined. (In particular, a "matrix inverse" cannot be the solution [38].)

 $y_i = [\mathcal{A}f]_i + \varepsilon_i,$ 

e,four,op,Aop

#### 6.1.4.1 Adjoint operator

A few operators related to  $\mathcal{A}$  will appear frequently. Throughout this chapter, we use the usual inner products on  $\mathcal{L}_2(\mathcal{D})$  and  $\mathbb{C}^{n_d}$ , for which the **adjoint** of  $\mathcal{A}$ , denoted  $\mathcal{A}^*$ , is given by

$$g = \mathcal{A}^* \boldsymbol{w} \iff g(\vec{\mathbf{x}}) = \sum_{i=1}^{n_d} w_i \, \phi_i(\vec{\mathbf{x}}) = \begin{cases} \sum_{i=1}^{n_d} w_i \, \mathrm{e}^{i 2\pi \vec{\nu}_i \cdot \vec{\mathbf{x}}}, & \vec{\mathbf{x}} \in \mathcal{D} \\ 0, & \text{otherwise.} \end{cases}$$
(6.1.10)

In Fourier transform terminology,  $\mathcal{A}$  is the **analysis** operator and  $\mathcal{A}^*$  is a kind of **synthesis** operator. To verify that  $\mathcal{A}^*$  is the adjoint of  $\mathcal{A}$ , note that

$$\langle f, \mathcal{A}^* \boldsymbol{v} \rangle = \int f(\vec{\mathbf{x}}) \left[ \sum_{i=1}^{n_d} v_i \, \phi_i(\vec{\mathbf{x}}) \right]^* d\vec{\mathbf{x}} = \sum_{i=1}^{n_d} v_i^* \left[ \int \phi_i^*(\vec{\mathbf{x}}) \, f(\vec{\mathbf{x}}) \, d\vec{\mathbf{x}} \right] = \sum_{i=1}^{n_d} \left[ \mathcal{A} \, f \right]_i v_i^* = \langle \mathcal{A} \, f, \, \boldsymbol{v} \rangle$$

#### 6.1.4.2 Frame operator

The linear operator  $\mathcal{A}^*\mathcal{A}$ , known as the **frame operator** in sampling theory [39–41], acts as follows:

$$\mathcal{A}^* \mathcal{A} f = \sum_{i=1}^{n_{\rm d}} \langle f, \phi_i \rangle \phi_i \,. \tag{6.1.11}$$

This operator is *almost* spatially shift invariant, as seen by the following argument:

$$[\boldsymbol{\mathcal{A}}^*\boldsymbol{\mathcal{A}}f](\vec{\mathbf{x}}) = \sum_{i=1}^{n_{d}} \phi_i(\vec{\mathbf{x}}) \langle f, \phi_i \rangle = \sum_{i=1}^{n_{d}} e^{i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}} \mathbb{I}_{\{\vec{\mathbf{x}} \in \mathcal{D}\}} \int_{\mathcal{D}} f(\vec{\mathbf{x}}') e^{-i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}'} \, \mathrm{d}\vec{\mathbf{x}}' = \int_{\mathcal{D}} f(\vec{\mathbf{x}}') \, h(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \, \mathrm{d}\vec{\mathbf{x}}',$$

where the **impulse response** of the operator  $\mathcal{A}^*\mathcal{A}$  is given by

$$h(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \triangleq \sum_{i=1}^{n_{d}} e^{i2\pi\vec{\nu}_{i} \cdot (\vec{\mathbf{x}} - \vec{\mathbf{x}}')} \mathbb{I}_{\{\vec{\mathbf{x}} \in \mathcal{D}\}} \mathbb{I}_{\{\vec{\mathbf{x}}' \in \mathcal{D}\}}.$$
(6.1.12)

If we had  $\mathcal{D} = \mathbb{R}^{n_d}$ , then this impulse response would be shift invariant, but in general it is not due to the support indicator functions. However, this operator is *locally* approximately shift invariant. At the center of the FOV, the **local** impulse response of  $\mathcal{A}^* \mathcal{A}$  is

$$h(\vec{\mathbf{x}},\vec{\mathbf{0}}) = (\boldsymbol{\mathcal{A}}^* \boldsymbol{\mathcal{A}} \delta)(\vec{\mathbf{x}}) = \sum_{i=1}^{n_d} e^{i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}} \mathbb{I}_{\{\vec{\mathbf{x}} \in \mathcal{D}\}},$$
(6.1.13)

assuming that  $\vec{0} \in \mathcal{D}$ . The corresponding local frequency response of  $\mathcal{A}^* \mathcal{A}$  is

$$H(\vec{\nu}) = \sum_{i=1}^{n_{\rm d}} S(\vec{\nu} - \vec{\nu}_i), \qquad (6.1.14)$$

In other words, the following "shift invariant" approximation can be useful:

$$\mathcal{A}^* \mathcal{A} \approx \mathcal{F}^{-1} \mathcal{D}(H(\vec{\nu})) \mathcal{F}, \qquad (6.1.15)$$

where  $\mathcal{F}$  is the  $\bar{d}$ -dimensional Fourier transform operator and  $\mathcal{D}(H(\vec{\nu}))$  is defined by

$$G = \mathcal{D}(H(\vec{\nu})) F \iff G(\vec{\nu}) = H(\vec{\nu}) F(\vec{\nu}), \qquad \forall \vec{\nu} \in \mathbb{R}^{n_{\rm d}}.$$

#### 6.1.4.3 Crosstalk matrix

The  $n_d \times n_d$  matrix  $\mathbf{K} \triangleq \mathcal{A}\mathcal{A}^*$  has been called the Fourier crosstalk matrix [42] and the point set matrix. This matrix would be diagonal if the frequency components  $\phi_i$  were orthogonal, such as would be the case if the frequency samples  $\vec{\nu}_i$  were equally spaced with separation appropriate for the domain  $\mathcal{D}$ . The elements of  $\mathcal{A}\mathcal{A}^*$  are given by

$$[\mathcal{A}\mathcal{A}^*]_{il} = \langle \phi_i, \phi_l \rangle = \int_{\mathcal{D}} e^{-i2\pi \vec{\nu}_i \cdot \vec{x}} e^{i2\pi \vec{\nu}_l \cdot \vec{x}} d\vec{x}$$

e.four.op.A\*A.fourier

$$= \int_{\mathcal{D}} e^{-i2\pi(\vec{\nu}_i - \vec{\nu}_l) \cdot \vec{x}} \, \mathrm{d}\vec{x} = S(\vec{\nu}_i - \vec{\nu}_l), \tag{6.1.16}$$

where  $S(\vec{\nu})$  was defined in (6.1.7).

It is easily shown (see Problem 6.7) that the functions  $\{\phi_i\}$  are linearly independent when the frequencies  $\vec{\nu}_i$  are distinct, and hence the crosstalk matrix is positive definite (though possibly poorly conditioned depending on the sampling pattern). For general sampling patterns  $\{\vec{\nu}_i\}$ , the crosstalk matrix is full, so it is impractical to store it explicitly for large  $n_d$ .

**Example 6.1.1** In the usual case where the FOV is square, i.e.,  $\mathcal{D} = [-\text{FOV}/2, \text{FOV}/2]^{\bar{d}}$ , the elements of the Fourier crosstalk matrix are samples of sinc functions:  $[\mathcal{AA}^*]_{il} = S(\vec{\nu}_i - \vec{\nu}_l)$ , where  $S(\vec{\nu}) = \text{FOV}^{\bar{d}} \operatorname{sinc}_{\bar{d}}(\text{FOV} \vec{\nu})$ . If the frequency sample locations are integer multiples of 1/FOV, then because of the positions of the zeros of a sinc function, the Fourier crosstalk matrix would be a scaled identity matrix:  $\mathcal{AA}^* = \text{FOV}^{\bar{d}}\mathbf{I}$ . Furthermore, in the 1D case with  $\nu_i = [i - (n_d + 1)/2]/\text{FOV}$ , the impulse response (6.1.12) of the frame operator is a Dirichlet kernel:

$$h(x, x') = \frac{\sin(n_{\mathrm{d}}\pi x/\mathrm{FOV})}{\sin(\pi x/\mathrm{FOV})} \mathbb{I}_{\{x \in \mathcal{D}\}} \mathbb{I}_{\{x' \in \mathcal{D}\}}$$

#### 6.1.4.4 Weighted crosstalk matrix

Define a "diagonal" image-domain weighting operator  $\mathcal{W} = \mathcal{D}(w(\vec{x}))$  by

$$g = \mathcal{W}f \iff g(\vec{\mathbf{x}}) = w(\vec{\mathbf{x}}) f(\vec{\mathbf{x}}), \quad \forall \vec{\mathbf{x}} \in \mathbb{R}^d.$$

Then generalizing (6.1.16) leads to the following weighted crosstalk matrix:

$$\left[\mathcal{AWA}^{*}\right]_{il} = \int_{\mathcal{D}} w(\vec{\mathbf{x}}) e^{-i2\pi(\vec{\nu}_{i} - \vec{\nu}_{l})} d\vec{\mathbf{x}} = \int s(\vec{\mathbf{x}}) w(\vec{\mathbf{x}}) e^{-i2\pi(\vec{\nu}_{i} - \vec{\nu}_{l})\vec{\mathbf{x}}} d\vec{\mathbf{x}} = (S * W)(\vec{\nu}_{i} - \vec{\nu}_{l}), \qquad (6.1.17)$$

where  $S(\vec{\nu})$  was defined in (6.1.7) and similarly  $W(\vec{\nu})$  is the  $\bar{d}$ -dimensional Fourier transform of  $w(\vec{x})$ .

## **6.2** Finite-series (discrete-discrete) methods (s,four,series)

Because we observe only a finite number of measurements, and because computers have finite memory (and display devices have finite pixels), it is natural to consider models for the object f that have a finite number  $n_p$  of unknown parameters. Such models will only approximate the true object, but they can be useful nevertheless. Linear models are the easiest to analyze and are hence the most common, so we focus on such models here.

### 6.2.1 Finite-series object model

For a finite-series approach, we first select some basis functions<sup>1</sup>  $\{b_j(\vec{x}) : j = 1, ..., n_p\}$ , and model the object f as follows<sup>2</sup>:

$$f(\vec{\mathbf{x}}) \approx \sum_{j=1}^{n_{\rm p}} x_j \, b_j(\vec{\mathbf{x}}) \,.$$
 (6.2.1)

After adopting such a model, the reconstruction problem simplifies to determining the vector of unknown coefficients  $\boldsymbol{x} = (x_1, \dots, x_{n_p})$  from the measurement vector  $\boldsymbol{y}$ . Defining the "synthesis" operator<sup>3</sup>  $\mathcal{B}_{\Box} : \mathbb{C}^{n_p} \to \mathcal{L}_2(\mathcal{D})$  by

$$[\mathcal{B}_{\Box} \boldsymbol{x}](\vec{\mathbf{x}}) = \sum_{j=1}^{n_{\mathrm{p}}} x_j \, b_j(\vec{\mathbf{x}}), \tag{6.2.2}$$

we can also write (6.2.1) as  $f \approx \mathcal{B}_{\Box} x$ .

<sup>&</sup>lt;sup>1</sup> Recall that the definition of a **basis** implies that the functions are **linearly independent** [43, p. 20]; this ensures that the series representation (6.2.1) has a unique set of coefficients.

<sup>&</sup>lt;sup>2</sup> In this expression,  $\vec{x} \in \mathbb{R}^{\bar{d}}$  denotes spatial coordinates, whereas  $\boldsymbol{x} = (x_1, \dots, x_{n_p}) \in \mathbb{C}^{n_p}$  denotes series coefficients.

<sup>&</sup>lt;sup>3</sup> The usual basis is a square pixel so the square subscript serves as a reminder, at least to me.

Under the approximation (6.2.1), the discrete data, continuous object model (6.1.9) simplifies to the following discrete-discrete model:

$$y = Ax + \varepsilon, \tag{6.2.3}$$

where the elements of the  $n_{\rm d} \times n_{\rm p}$  system matrix A are given by

$$a_{ij} = \int_{\mathcal{D}} b_j(\vec{\mathbf{x}}) e^{-i2\pi \vec{\nu}_i \cdot \vec{\mathbf{x}}} d\vec{\mathbf{x}} = \langle b_j, \phi_i \rangle.$$
(6.2.4)

One option for the basis functions is to use Dirac impulses:

$$b_j(\vec{\mathbf{x}}) = \delta(\vec{\mathbf{x}} - \vec{\mathbf{x}}_j),\tag{6.2.5}$$

where  $\vec{x}_j \in D$  denotes the location of the *j*th impulse. This choice is often used implicitly, and sometimes explicitly [41, 44]. For this choice, the elements of A are simply complex exponentials:  $a_{ij} = e^{-i2\pi \vec{\nu}_i \cdot \vec{x}_j}$ , and one can view A as a "discretization" of the operator A.

In general, we can write the matrix A concisely as follows

$$A=\mathcal{AB}_{\square}$$
 .

The relationships between the models are summarized as follows:

$$f \to [\mathcal{A}] \to \bar{\mathcal{y}}$$
 continuous to discrete mapping,  
 $x \to [\mathcal{B}_{\Box}] \to f$  finite-series expansion,  
 $x \to [\mathcal{A} = \mathcal{A} \mathcal{B}_{\Box}] \to \bar{\mathcal{y}}$  discrete-discrete mapping.

For the discrete-discrete model (6.2.3), there are a variety of possible methods for determining x, a few of which are described next.

#### 6.2.2 Weighted least-squares (WLS) solution

For some choices of frequency samples  $\{\vec{v}_i\}$  and basis functions  $\{b_j(\vec{x})\}$ , the matrix A has full column rank. For example, if  $\bar{d} = 1$  then having  $n_d \ge n_p$  (and distinct frequency samples) is sufficient to ensure A has rank  $n_p$  for equally-spaced bases [41, Lemma 1] [45]. For  $\bar{d} \ge 1$ , if  $n_d \ge 2n_p + \bar{d} - 1$ , then a generalization of Carathéodory's uniqueness result ensures that A has full column rank for equally-spaced bases [46]. In such cases, the matrix A'A will be positive definite and one could then consider a weighted least-squares (WLS) estimate of x:

$$\hat{x} = \arg\min_{x \in \mathbb{C}^{n_{\mathrm{p}}}} \|y - Ax\|_{W^{1/2}}^{2} = [A'WA]^{-1}A'Wy,$$
(6.2.6)

for some positive definite weighting matrix W. On the other hand, if  $n_d < n_p$ , then clearly A will be rank deficient so the LS criterion does not specify a unique solution. Efficient characterization of the invertibility of A'A is challenging [45], although some results are available for random sampling patterns [47].

Even in cases where generalizations of Carathéodory's result ensure uniqueness, stability is not ensured in general. If the frequency samples are spaced nonuniformly, then usually the matrix A'A will be ill conditioned. And if  $n_d < n_p$ , then A'A will be outright singular. So often the LS estimate (6.2.6) will be too noisy to be useful.

#### 6.2.3 Penalized weighted least-squares (PWLS) method

To control noise, one can use the series representation (6.2.1) to translate a continuous-space penalty function  $R_0(f)$  such as (6.3.6), defined in terms of the continuous-space object f, into a corresponding penalty function R(x), defined in terms of the coefficient vector x, as follows:

$$\mathsf{R}(\boldsymbol{x}) = \mathsf{R}_0(\boldsymbol{\mathcal{B}}_{\scriptscriptstyle \Box}\boldsymbol{x}) \,. \tag{6.2.7}$$

Alternatively, one can define a convenient roughness penalty function R(x) in terms of the coefficients x directly (see Chapter 2). In either case, we can then define an estimator  $\hat{x}$  as the minimizer of a PWLS cost function of the form

$$\Psi(\boldsymbol{x}) = \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}^{1/2}}^2 + \mathsf{R}(\boldsymbol{x}).$$
(6.2.8)

If the penalty function is quadratic, *i.e.*,  $R(x) = \frac{1}{2}x'Rx$  for some  $n_p \times n_p$ , Hermitian symmetric, positive-semidefinite matrix **R**, then the quadratically penalized WLS (**QPWLS**) estimate is

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \Psi(\boldsymbol{x}) = \left[\boldsymbol{A}' \boldsymbol{W} \boldsymbol{A} + \boldsymbol{\mathsf{R}}\right]^{-1} \boldsymbol{A}' \boldsymbol{W} \boldsymbol{y}. \tag{6.2.9}$$

Usually the null space of R is disjoint from the null space of A, so the above inverse exists. One would almost never compute  $\hat{x}$  using the expression (6.2.9). Instead, one computes  $\hat{x}$  by using an iteration like the **conjugate gradient** algorithm to minimize the cost function (6.2.8), as described in §6.2.9.

Using the series expansion (6.2.1), if desired we could define a continuous-space estimate  $\hat{f}$  in terms of the vector QPWLS estimate  $\hat{x}$  as follows:

$$\hat{f} = \mathcal{B}_{\Box} \hat{x} = \mathcal{B}_{\Box} [A'WA + R]^{-1} A'Wy.$$
(6.2.10)

This last step is unnecessary for practical purposes when we use square pixels as the basis functions and simply display the coefficients  $\hat{x}$  directly. But the expression (6.2.10) is useful for theoretical comparisons.

#### 6.2.4 Choosing the regularization parameter

As described in Chapter 1, a typical quadratic roughness penalty has the form  $R(x) = \beta \frac{1}{2} ||Cx||^2$ , for a "differencing" matrix C. A common concern with regularized methods like (6.2.10) is choosing the regularization parameter  $\beta$ . Based on (6.2.9) we can analyze the spatial resolution properties of  $\hat{f}$  easily:

$$\mathsf{E}\Big[\hat{f}\Big] = \mathcal{B}_{\scriptscriptstyle \Box}[A'WA + \beta C'C]^{-1}A'W\mathcal{A}f.$$

To examine the local impulse response at the center of the FOV, consider the case where  $f(\vec{x}) = \delta(\vec{x})$  in which case  $\mathcal{A} f = 1$ . Furthermore, in the usual case where W = I, we have

$$\mathsf{E}\Big[\hat{f}\Big] = \mathcal{B}_{\scriptscriptstyle \Box}[A'A + \beta C'C]^{-1}A'1.$$

Using FFTs, as described in Chapter 1, this local PSF is evaluated rapidly. One can then vary  $\beta$  and choose the value that yields the desired FWHM of the local PSF [48].

Alternatively, one can analyze spatial resolution properties using the discrete-discrete model (6.2.3) for the QPWLS estimator (6.2.9), for which

$$\mathsf{E}[\hat{\boldsymbol{x}}] = \left[\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A} + \boldsymbol{\beta}\boldsymbol{C}'\boldsymbol{C}\right]^{-1}\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\boldsymbol{x}.$$

Setting  $x = e_j$  leads to a discrete PSF, and one can choose  $\beta$  such that this PSF has a desired FWHM, *e.g.*, 1.5 pixels. MIRT See <code>qpwls\_psf.m</code>.

Another option is to use the automatic methods described in §2.5.

### 6.2.5 Equally-spaced basis functions

Usually the basis functions in (6.2.1) are chosen to be equally spaced translates of a pulse-like function such as a rectangle or triangle, although more complicated choices such as **prolate spheroidal wave functions** have also been used [49]. Specifically, usually we have

$$b_j(\vec{\mathbf{x}}) = b(\vec{\mathbf{x}} - \vec{\mathbf{x}}_j), \qquad j = 1, \dots, n_{\rm p},$$
(6.2.11)

where  $b(\vec{x})$  denotes the common function that is translated to  $\vec{x}_j$ , the center of the *j*th basis function.

For basis functions of the form (6.2.11), by the shift property of the Fourier transform, the elements of A in (6.2.4) are simply

$$a_{ij} = B(\vec{\nu}_i) \,\mathrm{e}^{-i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}_j} \,, \tag{6.2.12}$$

where  $B(\vec{\nu})$  is the  $\bar{d}$ -dimensional Fourier transform of  $b(\vec{x})$ . In other words, the system matrix A has the following form:

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{E},\tag{6.2.13}$$

where  $\boldsymbol{B}$  is a  $n_d \times n_d$  diagonal matrix:  $\boldsymbol{B} = \text{diag}\{B(\vec{\nu}_i)\}$ , and  $\boldsymbol{E} \in \mathbb{C}^{n_d \times n_p}$  has elements  $E_{ij} = e^{-i2\pi \vec{\nu}_i \cdot \vec{x}_j}$ . In MRI, the matrix  $\boldsymbol{E}$  is sometimes called the Fourier encoding matrix.

#### 6.2.6 Choice of basis functions

As described in Chapter 10, there are numerous possible choices of basis functions  $b(\vec{x})$  that have been used in various image reconstruction problems. In the specific context of reconstruction from Fourier samples, the following choices seem natural.

- Some papers explicitly or implicitly assume a Dirac impulse:  $b(\vec{x}) = \delta(\vec{x})$ . For this choice, B = I, which offers a slight simplification, but clearly the model is unrealistic physically.
- Another choice that leads to B = I is to a use sinc-like or jinc-like basis function whose support in the frequency domain tightly covers the set of sample locations {ν<sub>i</sub>}, *i.e.*, the convex hull thereof. In other words, choose b(x) such that B(ν<sub>i</sub>) = 1. In 1D, if ν<sub>max</sub> = max<sub>i</sub> |ν<sub>i</sub>|, then b(x) = 2ν<sub>max</sub> sinc(x2ν<sub>max</sub>) is one choice that satisfies this property. However, a limitation of this choice is that sinc and jinc functions are not space limited.
- Often we choose  $b(\vec{x})$  simply to be the indicator function corresponding to the desired pixel or voxel size. (This is particularly natural because images are usually viewed on computer displays having square pixels of (nearly) uniform luminance.) For  $\vec{d} = 1$ , a typical choice is the rectangular function:  $b(x) = \text{rect}(x/\Delta_x)$ , where  $\Delta_x$  is the spacing of the basis functions. Of course, the discontinuous nature of these basis functions is also unrealistic physically.

Clearly none of the above options is uniquely ideal. Whether there are other choices of basis functions, such as Kaiser-Bessel functions, that could be preferable is an *open problem*.

To understand the effect of the choice of  $b(\vec{x})$ , rewrite the QPWLS estimate (6.2.9) using (6.2.13):

$$\hat{oldsymbol{x}} = \left[oldsymbol{E}' ilde{oldsymbol{W}}oldsymbol{E} + oldsymbol{R}
ight]^{-1}oldsymbol{E}' ilde{oldsymbol{W}}oldsymbol{B}^{-1}oldsymbol{y},$$

where we define  $\hat{W} \triangleq B'WB$ . We can interpret this  $\hat{x}$  as the QPWLS estimate with modified weighting matrix  $\hat{W}$ , pre-scaled data  $B^{-1}y$ , and the Dirac impulse basis (6.2.5). In particular, if we were to choose the weighting matrix  $W = (BB')^{-1}$ , for which there would be little justification, then the QPWLS estimate simplifies as follows:

$$\hat{f} = \mathcal{B}_{\scriptscriptstyle \Box} [ \boldsymbol{E}' \boldsymbol{E} + \boldsymbol{R} ]^{-1} \boldsymbol{E}' \boldsymbol{B}^{-1} \boldsymbol{y}$$

However, usually B has smaller values for higher spatial frequencies, so such a weighting would probably overemphasize high-frequency noise. Usually W = I is used instead (see Problem 6.11).

#### 6.2.7 Uniform frequency samples case

Although the case of nonuniformly spaced frequency samples is more interesting, it can help clarify the ideas to consider the case of equally spaced frequencies. For example, in 1D we could have

$$\vec{\nu}_i = (i - (n_d + 1)/2) \Delta_{\nu}, \qquad i = 1, \dots, n_d,$$
(6.2.14)

where  $\triangle_{\nu}$  denotes the spatial frequency sample spacing. Similarly for  $\bar{d} > 1$ .

In this case, if  $n_d = n_p$  and  $\Delta_{\nu} = 1/(n_p \Delta_x)$ , then E is an orthogonal matrix, with  $E'E = n_p I$  and  $E^{-1} = \frac{1}{n_p}E'$ . Here, regularization is unnecessary because this case is perfectly conditioned, *i.e.*, the condition number of E'E is unity. Substituting R = 0 into (6.2.10) and simplifying yields

$$\hat{f} = \mathcal{B}_{\scriptscriptstyle \Box} \left( rac{1}{n_{
m p}} E' B^{-1} oldsymbol{y} 
ight).$$

This is essentially the conventional "inverse FFT" approach used for reconstruction from equally spaced frequency samples.

#### 6.2.8 Complex exponential basis

An alternative to the "shifted" basis (6.2.11) would be to use complex exponentials with equally-spaced frequencies. For example, in 1D we could use

$$b_{i}(x) = e^{i2\pi(j - (n_{\rm p} + 1)/2)\Delta_{\nu}x} s(x).$$
(6.2.15)

In fact, because the object f has a finite support D, it has an exact Fourier series representation akin to the form (6.2.15), but with an infinite number of terms. Further consideration of this choice is left as an exercise.

#### **6.2.9 Conjugate-gradient algorithm for practical PWLS estimation** (s,four,series,pcg)

Chapters 11, 14 and 15 describe algorithms for computing  $\hat{x}$  by finding the minimizer of cost functions like (6.2.8). However, unlike in tomography, here A is *not* a sparse matrix. In fact, A is usually too large to store explicitly; instead, all we can afford to store is the ingredients that define A in (6.2.13), namely the diagonal of B, the frequencies  $\{\vec{v}_i\}$ , and the spatial locations  $\{\vec{x}_j\}$ . Thus, many iterative optimization algorithms are inefficient for this application. A notable exception is the preconditioned **conjugate gradient** (PCG) algorithm of §11.8 and §14.6.2. The key step in any gradient-based descent algorithm such as PCG is computing the gradient of  $\Psi(x)$ , which has the form

$$\nabla \Psi(\boldsymbol{x}) = -\boldsymbol{A}' \boldsymbol{W}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}) + \nabla \mathsf{R}(\boldsymbol{x}).$$
(6.2.16)

The computational bottlenecks are computing the matrix-vector multiplication Ax and its transpose A'v, without storing A or A' explicitly.

Fortunately, there are efficient and very accurate algorithms for computing these matrix-vector multiplications by using nonuniform **fast Fourier transform** (**NUFFT**) approximations [18, 50]. Specifically, each multiplication by A or A' requires an over-sampled FFT and some simple interpolation operations. (This operation is akin to "reverse gridding," *e.g.*, [51–53].) See §6.6 for an overview. One can precompute and store the interpolation coefficients or compute them as needed [50]. Particularly efficient methods are available for gaussian interpolation kernels [54]. Using the **optimization transfer** techniques described in Chapter 12, nonquadratic regularization can also be included<sup>4</sup>. So PWLS estimators based on (6.2.8) are feasible for routine practical use.

MIRT See Gmri.m and Gnufft.m and mri\_example.m.

Kadah et al. proposed to compute the 1D DFT of each row of A and discard the small values, yielding an approximation of the form  $A \approx SQ_1^{-1}$  where S is a sparse matrix and  $Q_1$  denotes the 1D DFT [56]. Then one can iteratively update  $Q_1x$  using CG using sparse matrix operations rather than FFTs each iteration. However, precomputing M is expensive (although it only needs to be done once for a given trajectory). In contrast, the precomputation required by the Toeplitz approach described next is simple, and requires no approximations.

#### 6.2.10 Toeplitz embedding

s.four.series.toepl

Usually the weighting matrix is diagonal, *i.e.*,  $W = \text{diag}\{w_i\}$ , and in fact usually W = I. And usually the basis functions are equally spaced, as in (6.2.11). In these cases, the **Gram matrix** A'WA associated with the norm term in (6.2.8) is **block Toeplitz** with **Toeplitz blocks (BTTB**), and has elements

$$[\mathbf{A}'\mathbf{W}\mathbf{A}]_{kj} = \sum_{i=1}^{n_{\rm d}} w_i |B(\vec{\nu}_i)|^2 \,\mathrm{e}^{-i2\pi\vec{\nu}_i \cdot (\vec{\mathbf{x}}_j - \vec{\mathbf{x}}_k)} \,. \tag{6.2.17}$$

In 1D, when  $\bar{d} = 1$ , we have  $\vec{x}_j - \vec{x}_k = (j - k)\Delta$ , where  $\Delta$  denotes the basis function spacing, and A'WA is simply Toeplitz. For  $\bar{d} > 1$ , the nature of  $\vec{x}_j - \vec{x}_k$  induces the block Toeplitz property. In other words, the discrete-space Gram matrix  $A'WA : \mathbb{C}^{n_p} \to \mathbb{C}^{n_p}$  is *almost* shift invariant, just as §6.1.4 showed that the **frame operator**  $\mathcal{A}^*\mathcal{A}$  is almost shift invariant.

By defining the Toeplitz matrix T = A'WA and the "back-projected data" vector b = A'Wy, we can rewrite the gradient expression (6.2.16) as follows:

$$\nabla \Psi(\boldsymbol{x}) = \boldsymbol{T}\boldsymbol{x} - \boldsymbol{b} + \nabla \mathsf{R}(\boldsymbol{x}). \tag{6.2.18}$$

The elements of  $\boldsymbol{b} \in \mathbb{C}^{n_{\mathrm{P}}}$  are given by

$$b_j = [\mathbf{A}'\mathbf{W}\mathbf{y}]_j = \sum_{i=1}^{n_d} w_i y_i e^{-\imath 2\pi \vec{\nu}_i \cdot \vec{x}_j}, \qquad j = 1, \dots, n_p.$$

We can precompute **b** prior to iterating using an (adjoint) NUFFT operation [50]. (See §6.6. This calculation is similar to the gridding reconstruction method described in §6.4.3.5.) Each gradient calculation requires multiplying the  $n_{\rm p} \times n_{\rm p}$  block Toeplitz matrix **T** by the current guess of **x**. That operation can be performed efficiently by **embedding T** into a  $2^{\bar{d}}n_{\rm p} \times 2^{\bar{d}}n_{\rm p}$  block circulant matrix and applying a  $\bar{d}$ -dimensional FFT [41, 44, 57–61]. (See §6.8.) This is called the **ACT method** in the band-limited signal interpolation literature [41, 44], and it has also

e.four.series.grad.]

<sup>&</sup>lt;sup>4</sup> The quadratic surrogates of Chapter 12 require potential functions with bounded curvatures. Generalized gaussian potentials, which have unbounded curvatures, have also been proposed, requiring alternative minimization methods [55].

been applied to MR image reconstruction both with [62, 63] and without [64] an over-sampled FFT. It keeps getting re-discovered [65].

The first row of the circulant matrix is constructed by using  $2^{\bar{d}-1}$  (adjoint) NUFFT calls to evaluate columns of (6.2.17). When  $\bar{d} = 1$ , a single (adjoint) NUFFT will evaluate the first column of (6.2.17). (See §6.6 for the case  $\bar{d} = 2$ .) (There are also non-iterative, recursive methods for inverting Toeplitz matrices, *e.g.*, [66], but these are less efficient.)

Using the gradient expression (6.2.16) requires two NUFFT operations per iteration. Each NUFFT requires an over-sampled FFT and frequency-domain interpolations. In contrast, by using the Toeplitz approach (6.2.18), each iteration requires two (double-sized) FFT operations. No interpolations are needed except in the precomputing phase of building T and b. For an accurate NUFFT, usually we oversample the FFT by a factor of two (in each dimension). Thus, the NUFFT approach and the Toeplitz approach require exactly the same amount of FFT effort, but the NUFFT approach has the disadvantage of also requiring interpolations. The only apparent drawback of the Toeplitz approach (6.2.18) is that it "squares the condition number" of the problem so may be less numerically stable. However, in most applications the measurement noise will dominate the numerical noise, and to control measurement noise one will need to include suitable regularization which will also reduce the condition number.

MIRT See nufft\_toep.m for construction of T, and mri\_example.m for a comparison.

Further simplifications are possible if the regularizer is quadratic with a Hessian **R** that is also block Toeplitz, since then one can also incorporate **R** into T to reduce computation when finding the QPWLS estimate (6.2.9).

Excellent circulant preconditioners are available to accelerate the convergence rate of the CG algorithm for such Toeplitz problems [58, 67].

#### s, four, series, pixel 6.2.11 Effect of number of basis functions and pixel size (s, four, series, pixel)

A philosophical objection to the finite-series model (6.2.1) is that it requires the apparently subjective choice of basis functions. One might imagine that using "too many, too small" basis functions would lead to unstable results. For an *unregularized* approach like (6.2.6), increasing  $n_p$  will certainly cause instability. However, for a *regularized* approach like (6.2.10), the only "problem" with increasing  $n_p$  beyond a certain point is that computation time increases without any improvement in image quality.

#### MIRT See mri\_pixel\_size\_example.m.

To illustrate, we simulated a spiral sampling pattern for a 2D object with a 256mm FOV with  $\max \|\vec{v}_i\| = \frac{1}{8}$  mm<sup>-1</sup>. For the object  $f^{\text{true}}$  shown in Fig. 6.2.1, consisting of 2D rect functions and gaussian functions, we computed  $n_d = 64^2$  noiseless samples along the spiral using the analytical expression for its Fourier transform. We used the iterative PCG algorithm described in §6.2.9 to compute QPWLS estimates  $\hat{x}$  for rectangular basis functions with  $n_p = N^2$ , where  $N = 32, 64, \ldots, 512$ . (As N increases, the pixel size decreases according to  $\Delta_x = \text{FOV}/N$ .) Fig. 6.2.1 shows the true object  $\hat{f}$  and the reconstructed images  $\hat{x}$ . When  $n_p$  is too small, the estimates are undesirably "blocky," particularly for N = 32. However, as N increases, the images become indistinguishable, *i.e.*, even when  $n_p \gg n_d$ , the regularization stabilizes the estimates.

Fig. 6.2.2 shows the central horizontal profiles through these images. The profile for the case N = 32 is poor, whereas the other profiles are indistinguishable. This particular spiral sampling pattern was designed to be appropriate for a  $64 \times 64$  image. Using more pixels increases computation time but yields indistinguishable results.

We conjecture that  $\hat{f}_{\Delta} \rightarrow \hat{f}$  as  $\Delta \rightarrow 0$ , where  $\hat{f}_{\Delta}$  denotes the QPWLS estimate in (6.2.10) using a basis consisting of differentiable, localized functions such as cubic B-splines with support  $\Delta$ , and  $\hat{f}$  denotes the continuous-space QPWLS estimator of (6.3.7), assuming the penalty functions are related as in (6.2.7). A starting point for proving this would be Theorem 6 of [44].

#### , four, series, lin 6.2.12 General linear reconstructor (s, four, series, lin)

The preceding sections have focused on WLS and PWLS estimators, but certainly there are other possible reconstruction methods that start from the linear model (6.2.3). Any linear estimator for x given y has the form  $\hat{x} = Zy$  for some  $n_p \times n_d$  matrix Z. The question then becomes how to choose Z.

One possible criterion would be to consider the class of linear estimators that are unbiased under the discretediscrete model (6.2.3), *i.e.*, those for which ZA = I, and choose among those matrices the particular Z that minimizes the variance of each  $\hat{x}_j$ . By the Gauss-Markov theorem [68, p. 141], the optimal Z in this sense is the WLS estimator (6.2.6) (if it exists) with  $W = (\text{Cov}\{y\})^{-1}$ .



Figure 6.2.1: True object  $f^{\text{true}}$  and QPWLS estimates  $\hat{f}$  for  $n_{\text{p}} = N \times N$  pixel basis functions. The tick marks<sub>*i*</sub>are\_ni\_pixel\_size\_image in mm units.



fig\_mri\_pixel\_size\_profile

Figure 6.2.2: Central horizontal profiles through the images shown in Fig. 6.2.1.

An alternative approach is to examine the properties of  $\hat{x}$  in terms of the original discrete data, continuous object linear model (6.1.1), as follows:

$$\mathsf{E}[\hat{x}_j] = \sum_{i=1}^{n_{\rm d}} z_{ji} \,\mathsf{E}[y_i] = \sum_{i=1}^{n_{\rm d}} z_{ji} \int_{\mathcal{D}} f(\vec{x}) \,\phi_i^*(\vec{x}) \,\mathrm{d}\vec{x} = \int_{\mathcal{D}} f(\vec{x}) \,h_j^*(\vec{x}) \,\mathrm{d}\vec{x}, \qquad h_j(\vec{x}) \triangleq \sum_{i=1}^{n_{\rm d}} z_{ji}^* \,\phi_i(\vec{x}),$$

where  $\phi_i$  was defined in (6.1.4). The response function  $h_j(\vec{x})$  has been called the "voxel function" [69]. One could choose Z so that the response function  $h_j(\vec{x})$  "best matches" some *desired* response function  $h_j^{\text{des}}(\vec{x})$ . Because  $h_j(\cdot) = \mathcal{A}^* \boldsymbol{b}_j$ , where  $\boldsymbol{b}_j$  is the *j*th column of Z', one can show [69] by a least-squares criterion  $\min_{Z} \|h_j^{\text{des}}(\cdot) - h_j(\cdot)\|_{\mathcal{K}^{-1/2}}$ , for some image-domain weighting function  $\mathcal{K}^{-1}$ , that

$$oldsymbol{b}_j = \left[oldsymbol{\mathcal{A}}\mathcal{K}^{-1}oldsymbol{\mathcal{A}}^*
ight]^{-1}oldsymbol{\mathcal{A}}\mathcal{K}^{-1}\,h_j^{ ext{des}}(\cdot)$$
 .

Although this Z would lead to the best match of the desired voxel response function, the corresponding estimator  $\hat{x}$  will not have the minimum variance provided by the WLS estimator.

#### s, four, maxent 6.2.13 Maximum entropy formulations (s, four, maxent)

Image reconstruction methods based on **maximum entropy** principles have been explored in numerous imaging applications, including the problem of reconstruction from Fourier samples, *e.g.*, [70]. A complication that arises in this context is that often the object f may be complex, so conventional maximum entropy methods are inapplicable. Hoch et al. compare various definitions of **entropy** for complex-valued functions [71], and show that for cases where A is orthogonal, all of the versions reduce simply to **shrinkage** of the spectral components. Wang and Zhao [72] proposed an iterative method that uses a version of steepest descent to minimize a cost function of the form (6.2.8) where the regularizer R(x) involves terms of the form  $|x_j| \log |x_j|$  as described in [71]. Constable and Henkelman strongly critique maximum entropy methods in the context of MR reconstruction from partial k-space data [73]. On the other hand, the entropy of the *image gradient* has been used successfully for motion correction [74].

#### sparse 6.2.14 Sparse reconstruction (s,four,sparse)

s,four,minnorm

A contemporary research topic in signal processing is the problem of reconstructing a signal that is sparse in the frequency domain from random samples in the time domain [75–78].

The dual problem that is relevant to image reconstruction from Fourier samples is the problem of reconstructing an image that is sparse from random sample in the Fourier domain. [79–81]. In some applications such as angiography, the image itself is sparse and sparse reconstruction methods can be applied directly. Often it is more natural to consider the image to have a sparse representation in terms of some orthogonal basis such as wavelets, or to assume that the *gradients* of the image are sparse [82], *i.e.*, that that image is piecewise constant. This is a rapidly evolving field.

## 6.3 Discrete-data, continuous-object methods (s,four,dc)

If the results of §6.2.11 are unconvincing, then one could attempt to circumvent completely any subjectivity associated with the choice of basis functions  $\{b_j(\vec{x})\}$  in (6.2.1) by returning to the original discrete data, continuous object model (6.1.9). This section describes such methods [83].

#### 6.3.1 Minimum-norm least-squares methods (s,four,minnorm)

There are many functions f that exactly satisfy y = A f; one approach to selecting one of these possible solutions is to first choose a reference function  $f_0$ , and then to select the solution to y = A f that is closest (in some sense) to  $f_0$ , *i.e.*,

$$\min_{f \in \mathcal{L}_2(\mathcal{D})} \|f - f_0\|_{\mathcal{K}^{-1/2}} \text{ s.t. } \boldsymbol{y} = \boldsymbol{\mathcal{A}} f,$$
(6.3.1)

where  $\mathcal{K}$  denotes some user-selected, self-adjoint, positive-definite weighting operator on  $\mathcal{L}_2(\mathcal{D})$ , and  $\|\cdot\|$  denotes the usual norm on  $\mathcal{L}_2(\mathcal{D})$  corresponding to the inner product defined in (6.1.3). Usually  $\mathcal{K} = \mathcal{I}$ , where  $\mathcal{I}$  denotes the identity operator on  $\mathcal{L}_2(\mathcal{D})$ . The solution to this problem is [43, p. 65] [84] (*cf.* (26.9.2)):

$$\hat{f} = f_0 + \mathcal{K}^{1/2} (\mathcal{A}\mathcal{K}^{1/2})^\dagger (\boldsymbol{y} - \mathcal{A} f_0)$$

e,four,mnls,norm



Figure 6.3.1: Illustration of minimum-norm least-squares solution.

$$= f_0 + \mathcal{K}\mathcal{A}^* [\mathcal{A}\mathcal{K}\mathcal{A}^*]^{\dagger} (\boldsymbol{y} - \mathcal{A}f_0), \qquad (6.3.2)$$

where  $\mathcal{C}^{\dagger}$  denotes the **pseudo-inverse** of  $\mathcal{C}$ . Fig. 6.3.1 illustrates this solution graphically.

The  $n_{\rm d} \times n_{\rm d}$  Gram matrix

$$\tilde{K} \triangleq \mathcal{AKA}^* \tag{6.3.3}$$

is a weighted version of the Fourier crosstalk matrix in (6.1.16) akin to (6.1.17).

As discussed in §6.1.4, K is non-singular for distinct frequency samples, so the pseudo-inverse of K in (6.3.2) can be expressed as an inverse. Van de Walle et al. [85] investigated estimates of the form (6.3.2) for the "usual" case where  $f_0 = 0$  and  $\mathcal{K} = \mathcal{I}$ , *i.e.*,

$$\hat{f}_{\text{MNLS}} = \mathcal{A}^{\dagger} \boldsymbol{y} = \mathcal{A}^* [\mathcal{A} \mathcal{A}^*]^{-1} \boldsymbol{y}, \qquad (6.3.4)$$

noting that the inverse of the crosstalk matrix could be precomputed if one needed to determine estimates for many different data vectors y but for the same set of frequency sample locations  $\{\vec{v}_i\}$ . They emphasized that the **minimum-norm least-squares** (MNLS) solution (6.3.4) requires discretization only for image *display*; there is no discretization of the object f in the problem formulation, unlike a series expansion like (6.2.1).

However, the distinction is smaller than one might see at first. If we choose the object basis functions in (6.2.1) to be the complex exponentials in (6.1.4), *i.e.*, we choose  $n_p = n_d$  and  $b_j(\vec{x}) = \phi_j(\vec{x})$ , or equivalently we choose  $\mathcal{B}_{\Box} = \mathcal{A}^*$ , then the QPWLS solution (6.2.10) with W = I and R = 0 simplifies exactly to the MNLS solution (6.3.4). So (6.3.4) is simply a special case of the more general series formulation (6.2.1).

The basis choice  $\mathcal{B}_{\Box} = \mathcal{A}^*$  is sometimes described as **natural pixels** [86–88]; a discomforting property of this choice is that it depends on the set of frequency samples; changing those samples will change the object basis. Furthermore, although the choice of object basis functions in (6.2.1) is somewhat subjective, so too is the choice of reference image  $f_0$  in (6.3.2).

In summary, the MNLS solution has been suggested to be somehow more "pure" than finite-series methods, but the bottom line is that there are infinitely many exact "solutions" to y = A f, and "even more" approximate solutions. All criteria for singling out a particular estimate  $\hat{f}$  involve subjective preferences about the expected characteristics of f.

The bias of the MNLS method (6.3.2) is

$$\begin{split} \mathsf{E}\Big[\hat{f}\Big] - f &= f_0 + \mathcal{K}\mathcal{A}^*[\mathcal{A}\mathcal{K}\mathcal{A}^*]^{\dagger}(\mathcal{A}f - \mathcal{A}f_0) - f \\ &= \left(\mathcal{I} - \mathcal{K}\mathcal{A}^*[\mathcal{A}\mathcal{K}\mathcal{A}^*]^{\dagger}\mathcal{A}\right)(f_0 - f) \\ &= \mathcal{K}^{1/2}\left(\mathcal{I} - \mathcal{K}^{1/2}\mathcal{A}^*[\mathcal{A}\mathcal{K}\mathcal{A}^*]^{\dagger}\mathcal{A}\mathcal{K}^{1/2}\right)\mathcal{K}^{-1/2}(f_0 - f) \\ &= \mathcal{K}^{1/2}\mathcal{P}_{\mathcal{K}^{1/2}\mathcal{A}^*}^{\perp}\mathcal{K}^{-1/2}(f_0 - f), \end{split}$$

where  $\mathcal{P}_{\mathcal{C}}^{\perp}$  denotes the orthogonal projection onto the subspace perpendicular to the range of  $\mathcal{C}$ . Unsurprisingly, the bias depends on the choice of reference image  $f_0$ .

#### 6.3.2 Penalized least-squares for discrete data, continuous object model (s,four,pls)

In many of the applications where (6.1.9) applies, the noise has (at least approximately) a gaussian distribution. So by analogy with the regularized least-squares approach described in §1.8 and §4.3.3, a natural approach to estimating f

fig,four,minnor

would be to find the minimizer  $\hat{f}$  of a **penalized weighted-least squares** (**PWLS**) cost function of the form

$$\hat{f} = \underset{f \in \mathcal{L}_{2}(\mathcal{D})}{\arg\min} \Psi(f), \qquad \Psi(f) = \frac{1}{2} \| \boldsymbol{y} - \mathcal{A}f \|_{\boldsymbol{W}^{1/2}}^{2} + \mathsf{R}_{0}(f),$$
(6.3.5)

where  $W \in \mathbb{C}^{n_d \times n_d}$  is a user-selected, positive-definite weighting matrix. (By the Gauss-Markov theorem [68, p. 141], ideally W would be the matrix inverse of the covariance of y.) The functional  $R_0(f)$  denotes a continuous-space roughness penalty function such as those described in §2.4. Sophisticated roughness penalties based on methods such as level sets have been proposed, *e.g.*, [89]. For simplicity of analysis, consider the following first-order, rotationally-invariant quadratic roughness penalty

$$\mathsf{R}_{0}(f) = \int_{\mathcal{D}} \frac{1}{2} \left\| \nabla f \right\|^{2} \mathrm{d}\vec{\mathbf{x}} = \frac{1}{2} \left\| \mathcal{C} f \right\|^{2} = \frac{1}{2} \left\langle \mathcal{C} f, \mathcal{C} f \right\rangle = \frac{1}{2} \left\langle \mathcal{C}^{*} \mathcal{C} f, f \right\rangle, \tag{6.3.6}$$

where C f yields the gradient field of f. The first norm above is the *d*-space Euclidean norm whereas the second norm is the "usual" norm for  $\bar{d}$ -space gradient fields over  $\mathcal{L}_2(\mathcal{D})$ , *e.g.*, for  $\bar{d} = 2$ :

$$\left\| \mathcal{C} f \right\|^{2} = \int_{\mathcal{D}} \left| \frac{\partial}{\partial x_{1}} f(\vec{\mathbf{x}}) \right|^{2} + \left| \frac{\partial}{\partial x_{2}} f(\vec{\mathbf{x}}) \right|^{2} d\vec{\mathbf{x}}.$$

(Strictly speaking this is a **semi-norm**, a secondary technical detail.)

For such a quadratic regularizer, any minimizer  $\hat{f}$  of the cost function (6.3.5) must satisfy the following **normal** equations [43, p. 160]:

$$\left[\boldsymbol{\mathcal{A}}^*\boldsymbol{W}\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{R}}_0\right]\hat{f} = \boldsymbol{\mathcal{A}}^*\boldsymbol{W}\boldsymbol{y},\tag{6.3.7}$$

where  $\mathcal{R}_0 \triangleq \mathcal{C}^*\mathcal{C}$ . Usually the null space of  $\mathcal{R}_0$  consists only of functions that are constants over  $\mathcal{D}$ , and such functions are not in the null space of  $\mathcal{A}$  since the DC value  $\vec{\nu} = \vec{0}$  is usually available. So the operator  $[\mathcal{A}^*W\mathcal{A} + \mathcal{R}_0]$  is usually invertible (in principle). Unfortunately, unlike in the tomography problem described in §4.3.3, here the operator  $\mathcal{A}^*\mathcal{A}$  is *never* exactly shift invariant when the number of samples  $n_d$  is finite, so exact Fourier methods are inapplicable<sup>5</sup>, except perhaps for equally spaced frequencies (*cf.* Problem 6.9). I am unaware of any practical numerical procedures for solving (6.3.7), although some ideas are sketched in [90, 91].

If  $\mathcal{R}$  is approximately shift invariant with frequency response  $R(\vec{\nu})$ , then by (6.1.15), for  $\mathbf{W} = \mathbf{I}$  an approximate solution to the above normal equations is

$$\hat{f} \approx \boldsymbol{\mathcal{F}}^{-1} \, \boldsymbol{\mathcal{D}}\left(\frac{1}{H(\vec{\nu}) + R(\vec{\nu})}\right) \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{A}}^* \boldsymbol{y} = \boldsymbol{\mathcal{F}}^{-1} \, \boldsymbol{\mathcal{D}}\left(\frac{1}{H(\vec{\nu}) + R(\vec{\nu})}\right) \sum_{i=1}^{n_{\rm cl}} y_i \, S(\vec{\nu} - \vec{\nu}_i),$$

where  $H(\vec{\nu})$  was defined in (6.1.14). This approximation could be implemented easily using gridding and FFTs. The evaluation of such a method is an interesting open problem.

Although the formulations (6.1.9) and (6.3.5) allow f to be any continuous-space function, the solution (6.3.7) turns out to lie in a finite-dimensional subspace of  $\mathcal{L}_2(\mathcal{D})$ . In particular,

$$\hat{f}(\vec{x}) = \sum_{i=1}^{n_{\rm d}} g_i(\vec{x}) y_i,$$
(6.3.8)

where each  $g_i$  satisfies

$$\left[\boldsymbol{\mathcal{A}}^*\boldsymbol{W}\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{R}}_0\right]g_i=\boldsymbol{\mathcal{A}}^*\boldsymbol{W}\boldsymbol{e}_i,\qquad i=1,\ldots,n_{\rm d},$$

and where  $e_i$  denotes the *i*th unit vector in  $\mathbb{R}^{n_d}$ . Even though (6.3.5) is a nonparametric, continuous-space formulation, we see that it leads to a solution (6.3.8) that is a finite linear combination of certain functions  $\{g_i(\vec{x})\}$ . (These basis functions are related to the "equivalent kernels" as analyzed in the context of nonparametric regression [92].) The finite-series methods described in §6.2 could be considered to be generalizations that allow "basis functions" other than the above  $g_i$  functions.

four,pls,R0(f)

e,four,pls,normal

<sup>&</sup>lt;sup>5</sup> If the FOV is  $\mathcal{D} = \mathbb{R}^{n_d}$  and W is diagonal, then the Gram operator  $\mathcal{A}^* W \mathcal{A}$  is shift-invariant with impulse response  $h(\vec{x}) = \sum_{i=1}^{n_d} w_i e^{-i2\pi \vec{\nu}_i \cdot \vec{x}}$  and frequency response  $H(\vec{\nu}) = \sum_{i=1}^{n_d} w_i \delta(\vec{\nu} - \vec{\nu}_i)$ . But for any realistic bounded domain  $\mathcal{D}$ , the Gram operator is shift variant.

### s, four, dc, mnls 6.3.3 Relation between MNLS and QPWLS (s, four, dc, mnls)

The MNLS solution (6.3.4) is a special case of the continuous-space QPWLS solution of (6.3.7). Taking W = I and  $\mathcal{R} = \varepsilon \mathcal{K}^{-1}$  in (6.3.7) and applying the **push through identity** (26.1.10) leads to the QPWLS estimator

$$\begin{split} \hat{f}_{\varepsilon} &= \left[ \boldsymbol{\mathcal{A}}^* \boldsymbol{\mathcal{A}} + \varepsilon \boldsymbol{\mathcal{K}}^{-1} \right]^{-1} \boldsymbol{\mathcal{A}}^* \boldsymbol{y} \\ &= \boldsymbol{\mathcal{K}} \boldsymbol{\mathcal{A}}^* \left[ \boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{K}} \boldsymbol{\mathcal{A}}^* + \varepsilon \boldsymbol{I} \right]^{-1} \boldsymbol{y} \\ &\to \boldsymbol{\mathcal{K}} \boldsymbol{\mathcal{A}}^* [\boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{K}} \boldsymbol{\mathcal{A}}^*]^{\dagger} \boldsymbol{y} = \hat{f}_{\text{MNLS}}, \text{ as } \varepsilon \to 0. \end{split}$$

## 6.4 Continuous-continuous methods (s,four,cc)

Thus far we have described solution methods that are based on problem formulations that acknowledge explicitly the fact that the available data  $\boldsymbol{y}$  is discrete. Interestingly, the most common methods used for MR image reconstruction are based on solutions that are first formulated using a continuous-continuous model, and then the "harsh reality" that the data is discrete is incorporated *after* finding an analytical solution. The continuous model is simply the ordinary Fourier transform expression in (6.1.2). If we had available  $F(\vec{v})$  for all  $\vec{v} \in \mathbb{R}^{\bar{d}}$ , then it would be trivial to "derive" an analytical solution; we would simply use the inverse Fourier transform:

$$f(\vec{x}) = \int_{\mathbb{R}^d} F(\vec{\nu}) e^{i2\pi\vec{\nu}\cdot\vec{x}} \,\mathrm{d}\vec{\nu} \,.$$
(6.4.1)

This expression is the analog of the FBP or BPF methods for tomographic reconstruction; if we had  $F(\cdot)$ , we could compute f easily. Many papers on this problem begin with this "solution," even though the continuum of data  $\{F(\vec{\nu})\}$  is never available in practice. So (6.4.1) is wishful thinking.

To estimate f using (6.4.1) as a starting point, one must discretize the integral. (Of course, as soon as it is discretized, the method is no longer **continuous-continuous** so perhaps the section heading is a misnomer.) There are two general categories of discretization methods. One approach is to discretize (6.4.1) using the nonuniform sample locations  $\{\vec{v}_i\}$ . §6.4.1 describes the resulting **conjugate phase** reconstruction method. Alternatively, one can use a uniform discretization over  $\vec{v}$  by applying an frequency domain interpolation step, as described in §6.4.3. A particularly popular version of frequency domain interpolation is called **gridding**, as described in §6.4.3.5.

### four, conj 6.4.1 Conjugate phase method (s,four,conj)

To estimate f using (6.4.1) as a starting point, one approach is to discretize the integral using the nonuniform sample locations  $\{\vec{v}_i\}$ . Consider the following weighted summation estimator

$$\hat{f}(\vec{\mathbf{x}}) = \sum_{i=1}^{n_{\rm d}} y_i \, \mathrm{e}^{i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}} \, w_i \, s(\vec{\mathbf{x}}) \approx \sum_{i=1}^{n_{\rm d}} F(\vec{\nu}_i) \, \mathrm{e}^{i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}} \, w_i \, s(\vec{\mathbf{x}}), \tag{6.4.2}$$

where  $\{w_i\}$  denotes sampling density compensation factors that account for the nonuniform sampling density of the  $\vec{\nu}_i$  locations in  $\mathbb{R}^{\bar{d}}$ , and the spatial support  $s(\vec{x})$  was defined in (6.1.7). In MRI, this estimator is known as the conjugate phase reconstruction method. If  $f(\vec{x})$  (and  $s(\vec{x})$ ) are unitless and  $\vec{x}$  has units cm, then the units of  $F(\vec{\nu})$  are  $cm^{\bar{d}}$  and hence the units of  $w_i$  must be  $1/cm^{\bar{d}}$ .

Defining the sampling density compensation vector  $\boldsymbol{w} = (w_1, \dots, w_{n_d})$ , in operator notation the conjugate phase estimator is

$$\hat{f} = \mathcal{Z} D(w) y, \tag{6.4.3}$$

where  $D(w) = \text{diag}\{w_i\}$  and  $\mathcal{Z} : \mathbb{C}^{n_d} \to \mathcal{L}_2(\mathcal{D})$ . To match (6.4.2), the natural choice for  $\mathcal{Z}$  is

$$\mathcal{Z} = \mathcal{A}^*, \tag{6.4.4}$$

where  $\mathcal{A}^*$  was defined in (6.1.10). However, throughout this section we use the notation  $\mathcal{Z}$  rather than  $\mathcal{A}^*$  because approximations to  $\mathcal{A}^*$  are often used in practice. In MRI, an unweighted version of this estimator was first proposed by Macovski [93]. A year later the desirability of including weights was realized [94]. The weighted version has been called the weighted correlation method [95]. It is somewhat analogous to the FBP reconstruction method for tomography, considering  $\mathcal{Z}$  to be a kind of "backprojection" that maps frequency-space values back into complex exponential functions in object space. e, four, wish

In practice one evaluates  $f(\vec{x})$  at only a finite set of sample locations, *e.g.*,  $\vec{x} = n \Delta_x$ , n = -N/2, ..., N/2 - 1 in 1D. At these locations, the estimator (6.4.2) evaluates to:

$$\hat{f}(n\Delta_{\mathbf{x}}) = s(n\Delta_{\mathbf{x}}) \sum_{i=1}^{n_{\mathbf{d}}} w_i y_i \, \mathrm{e}^{i(2\pi\nu_i\Delta_{\mathbf{x}})n} \,. \tag{6.4.5}$$

The summation in (6.4.5) is the **adjoint** of the DSFT operation, and can be computed approximately using a NUFFT (see §6.6). Of course (6.4.5) maps a finite input vector y into a finite set of image samples, so one might call it a **discrete-discrete** method, but its origins are the **continuous-continuous** "solution" (6.4.1).

#### MIRT Gdsft' or Gnufft'

Comparing the conjugate phase estimator (6.4.3) to the MNLS solution (6.3.2) for the case  $f_0 = 0$ , we see that if we were to choose  $D = [\mathcal{AKA}^*]^{\dagger}$  and  $\mathcal{Z} = \mathcal{KA}^*$ , then the two estimators would be identical. However, in practice the conjugate phase solution (6.4.3) is usually, if not always, implemented with a *diagonal* matrix D(w), because a diagonal requires far less computation than the inverse of the Fourier cross-talk matrix.

The estimator (6.4.3) involves two nontrivial issues, both of which have been explored at length in the literature. One issue is computing efficiently the product Z times D(w)y. The usual approach, when  $Z = A^*$ , is called **gridding**, as discussed in §6.4.3.5. Alternative approximations have also been explored such as look-up tables [96] and equal phase lines [97]. Efficient methods have been developed for evaluating (6.4.2) using NUFFT tools [54, 98–107] (see §6.7). In essence, the computational issue is well understood. The next section focuses on the more vexing issue of choosing the density compensation factors w.

### s, four, dens 6.4.2 Sampling density compensation (s, four, dens)

If the conjugate phase estimator (6.4.3) were used with D = I, then spatial frequencies that are near higher sampling densities would be over-emphasized. For example, spiral and radial sampling patterns used in MRI have high sampling densities near  $\vec{\nu} = \vec{0}$ , so low spatial frequencies would be over-emphasized. So it is important to choose w carefully. Many methods have been proposed for choosing w. Many of the proposals have been described in the context of gridding-based reconstruction. We consider the general estimator (6.4.3), keeping Z distinct from  $A^*$  where possible.

#### ,four,dens,ld,cart

**Example 6.4.1** As a concrete example, consider the 1D case with equally spaced frequency samples:  $\nu_k = \frac{k}{\alpha FOV}$  for  $k \in \mathbb{Z}$ , where  $\alpha \ge 1$  is an "over-sampling" factor. Since f(x) is space-limited, by the sampling theorem we can reconstruct  $F(\vec{\nu})$  from its samples by sinc interpolation as follows:

$$F(\nu) = \sum_{k=-\infty}^{\infty} \operatorname{sinc}(\nu \alpha \text{FOV} - k) F\left(\frac{k}{\alpha \text{FOV}}\right)$$
$$\implies f(x) = \frac{1}{\alpha \text{FOV}} \operatorname{rect}\left(\frac{x}{\alpha \text{FOV}}\right) \sum_{k=-\infty}^{\infty} e^{i2\pi \frac{k}{\alpha \text{FOV}}x} F\left(\frac{k}{\alpha \text{FOV}}\right).$$

For this example, the proper choice for w is indisputably  $w_i = \frac{1}{\alpha FOV}$ , i.e.,  $w_i$  should be inversely proportional to the FOV and to the over-sampling factor. It is easy to generalize this example to higher dimensions.

For nonuniform sampling patterns, the best choice of w is less apparent. Methods for choosing w can be categorized as either heuristic or as being based on some optimality criterion.

#### 6.4.2.1 Noise considerations (s,four,dens,noise)

=

The covariance of the **conjugate phase** estimator (6.4.3) is

$$\operatorname{Cov}_{\boldsymbol{w}}\left\{\hat{f}
ight\} = \boldsymbol{\mathcal{Z}} \, \boldsymbol{D}(\boldsymbol{w}) \operatorname{Cov}\left\{\boldsymbol{y}
ight\} \boldsymbol{D}^{*}(\boldsymbol{w}) \, \boldsymbol{\mathcal{Z}}^{*},$$

where  $Cov{y} = \sigma^2 I$  for the usual additive white gaussian noise model. It is clear from this expression that larger values of the DCF elements  $w_i$  can increase the noise covariance.

For the estimator (6.4.2), the variance of the reconstructed image at any given spatial location is given by

$$\mathsf{Var}\Big\{\hat{f}(\vec{\mathbf{x}})\Big\} = s^2(\vec{\mathbf{x}})\,\sigma^2\sum_{i=1}^{n_{\mathrm{d}}}w_i^2 \approx \int w^2(\vec{k})d(\vec{k})\,\mathrm{d}\vec{k},$$

where  $d(\vec{k})$  denotes the sampling density at k-space location  $\vec{k}$ . Comparing this to iterative reconstruction methods is an interesting open problem.

One natural scalar measure of the "overall" noise in  $\hat{f}$  is the following weighted total noise variance:

$$\sigma_{\text{total}}^{2}(\boldsymbol{w}) \triangleq \text{trace}\Big\{\boldsymbol{\mathcal{K}}^{-1/2} \operatorname{Cov}_{\boldsymbol{w}}\Big\{\hat{f}\Big\} \boldsymbol{\mathcal{K}}^{-1/2}\Big\} = \sigma^{2} \operatorname{trace}\{\boldsymbol{D}(\boldsymbol{w}) \boldsymbol{D}^{*}(\boldsymbol{w}) \boldsymbol{J}\} = \sigma^{2} \sum_{i=1}^{n_{\text{d}}} |w_{i}|^{2} J_{ii} = \sigma^{2} \operatorname{trace}\{\boldsymbol{D}(\boldsymbol{w}) \boldsymbol{D}^{*}(\boldsymbol{w}) \boldsymbol{J}\} = \sigma^{2} \sum_{i=1}^{n_{\text{d}}} |w_{i}|^{2} J_{ii} = \sigma^{2} \operatorname{trace}\{\boldsymbol{D}(\boldsymbol{w}) \boldsymbol{D}^{*}(\boldsymbol{w}) \boldsymbol{J}\} = \sigma^{2} \sum_{i=1}^{n_{\text{d}}} |w_{i}|^{2} J_{ii} = \sigma^{2} \operatorname{trace}\{\boldsymbol{D}(\boldsymbol{w}) \boldsymbol{D}^{*}(\boldsymbol{w}) \boldsymbol{J}\} = \sigma^{2} \sum_{i=1}^{n_{\text{d}}} |w_{i}|^{2} J_{ii} = \sigma^{2} \operatorname{trace}\{\boldsymbol{D}(\boldsymbol{w}) \boldsymbol{D}^{*}(\boldsymbol{w}) \boldsymbol{J}\}$$

where  $\mathcal{K}$  is user-selected weighting and we define the following weighted relative of the crosstalk matrix in (6.3.3):

$$\boldsymbol{J} \triangleq \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}^{-1} \boldsymbol{\mathcal{Z}}.$$
 (6.4.6)

A natural choice for the weighting  $\mathcal{K}^{-1}$  is  $\mathcal{K}^{-1} = \mathcal{D}(c(\vec{x}))$ , where  $c(\vec{x})$  denotes an approximation to  $s(\vec{x})$  defined below in (6.4.53), and usually  $\mathcal{Z} = \mathcal{A}^*$ , in which case the diagonal elements of J are

$$J_{ii} = [\boldsymbol{\mathcal{Z}}^* \, \boldsymbol{\mathcal{D}}(c(\vec{\mathbf{x}})) \, \boldsymbol{\mathcal{Z}}]_{ii} = [\boldsymbol{\mathcal{A}} \, \boldsymbol{\mathcal{D}}(c(\vec{\mathbf{x}})) \, \boldsymbol{\mathcal{A}}^*]_{ii} = \int_{\mathcal{D}} c(\vec{\mathbf{x}}) \, \mathrm{d}\vec{\mathbf{x}},$$

which is a constant. In this typical case, the total noise variance is proportional to the norm of the DCF vector:

$$\sigma_{\text{total}}^{2}(\boldsymbol{w}) = \sigma^{2} \left( \int_{\mathcal{D}} c(\vec{x}) \, \mathrm{d}\vec{x} \right) \|\boldsymbol{w}\|_{2}^{2}.$$
(6.4.7)

In light of this analysis, it is natural to consider the norm  $\|\boldsymbol{w}\|$  of the DCF vector for DCF design, either by using a constraint or by using a regularizer of the form  $\beta \|\boldsymbol{w}\|_2^2$ . See §6.4.2.4.5.

#### s, heur 6.4.2.2 Heuristic methods (s, four, dens, heur)

Several heuristic methods for DCF design have been proposed, and many of these remain quite popular due to their simplicity.

**6.4.2.2.1** Jacobian determinant Some sampling patterns can be treated as a continuous and invertible mapping of uniformly spaced samples. For such patterns, the Jacobian determinant of the transformation (or approximations thereof) at each sample  $\vec{v}_i$  has been used to define w [23, 108–115]. This approach is computationally efficient compared to most of the alternatives below. However, it is specific to special sampling patterns, such as spirals, and it is difficult to accommodate self-crossing sampling trajectories or other departures from the ideal analytical formula for the sampling pattern [113]. This method also does not seem to account fully for the fact that  $n_d$  is finite. Nevertheless, it is particularly popular for certain sampling patterns such as spirals.

**6.4.2.2.2** Voronoi cell volume A simple approach that is applicable to arbitrary sampling patterns is to let  $w_i$  be the volume of the Voronoi cell around frequency sample  $\vec{v}_i$  [1, 116]. Fig. 6.4.1 shows a 2D example. Combinations of Voronoi cell volume and Jacobian determinants have also been suggested [117, 118].

Sample points at the outer boundaries of the sampling pattern (*i.e.*, points  $\vec{\nu}_i$  that are outside the convex hull of the other points) have Voronoi cells with infinite area. Thus one must somehow choose reasonable finite values for  $w_i$  for those locations.

MIRT voronoin, convhulln, mri\_density\_comp

**6.4.2.3 Cell counting** A simpler alternative to Voronoi cell volume is to partition frequency space into rectangular "Nyquist" cells (having dimensions corresponding to the reciprocals of the dimensions of the field of view  $\mathcal{D}$ ), and to let  $w_i$  be the reciprocal of the number of frequency samples in the cell that contains  $\vec{\nu}_i$ . This is called the **cell counting** method [44]. Presumably one must make adjustments for samples that lie on cell boundaries. Ignoring such corrections, the 1D version is:  $\mathcal{I}_k \triangleq [(k - 1/2)/\text{FOV}, (k + 1/2)/\text{FOV}), N_k \triangleq \sum_{i=1}^{n_d} \mathbb{I}_{\{\nu_i \in \mathcal{I}_k\}}, w_i = \sum_k \frac{1}{N_k} \mathbb{I}_{\{\nu_i \in \mathcal{I}_k\}}$ .

**6.4.2.2.4** Sinc overlap density The sinc function has the following "sampling density" property:

$$0 < \Delta < 2 \Longrightarrow \sum_{k=-\infty}^{\infty} \operatorname{sinc}(k\Delta) = \frac{1}{\Delta}.$$
(6.4.8)



Figure 6.4.1: Illustration of nonuniform frequency sampling pattern (the dots along the spiral) and the corresponding Voronoi cells. Areas of such cells have been used for sampling density compensation.

Since  $S(\vec{\nu})$  in (6.1.7) is usually a sinc-like function, this property suggests the following choice of sampling density compensation:

$$w_{i} = \frac{1}{\sum_{l=1}^{n_{d}} S(\vec{\nu}_{i} - \vec{\nu}_{l})} = \frac{1}{\sum_{l=1}^{n_{d}} [\mathcal{A}\mathcal{A}^{*}]_{il}} = \frac{1}{[\mathcal{A}\mathcal{A}^{*}\mathbf{1}]_{i}},$$
(6.4.9)

or equivalently

$$\boldsymbol{w} = \boldsymbol{1} \oslash (\mathcal{A}\mathcal{A}^*\boldsymbol{1}), \tag{6.4.10}$$

where  $\oslash$  denotes element-wise division and the  $n_d \times n_d$  crosstalk matrix  $\mathcal{AA}^*$  was defined in (6.1.16). This definition of w is ideal in the (uncommon) case of equally spaced frequency samples whose density exceeds the Nyquist spacing corresponding to  $\mathcal{D}$ .

**6.4.2.2.5** Jackson's area density In convolutional gridding methods such as [16, 119, 120] (see §6.4.3.5), the traditional choice for sampling density compensation is [16, eqn. (8)]:

$$w_i = \frac{1}{\sum_{l=1}^{n_d} C(\vec{\nu}_i - \vec{\nu}_l)}.$$
(6.4.11)

The denominator of (6.4.11) has been called the **area density**, where  $C(\vec{\nu})$  denotes the gridding kernel (such as a Kaiser-Bessel function); see (6.4.52). Presumably this choice is motivated in part by (6.4.9), because gridding kernels are finite-support approximations to  $S(\vec{\nu})$ . The other heuristic methods for sampling density compensation depend only on the sample locations  $\{\vec{\nu}_i\}$ , thereby avoiding the question of how to best choose  $C(\vec{\nu})$  in (6.4.11).

In light of (6.1.17), Jackson's area density can also be written *approximately* in terms of a **weighted crosstalk matrix** as follows:

$$\boldsymbol{w} \approx \mathbf{1} \oslash (\mathcal{A}\mathcal{D}(c(\vec{\mathrm{x}})) \mathcal{A}^* \mathbf{1}) = \mathbf{1} \oslash \left( \tilde{K} \mathbf{1} \right),$$
 (6.4.12)

where  $c(\vec{x}) \stackrel{\text{FT}}{\longleftrightarrow} C(\vec{\nu})$ , and where  $\tilde{K}$  was defined in (6.3.3) with  $\mathcal{K} = \mathcal{D}(c(\vec{x}))$  here. The approximation is valid provided  $c(\vec{x}) s(\vec{x}) \approx s(\vec{x})$ . This never holds exactly because  $C(\vec{\nu})$  has finite support for gridding, so  $c(\vec{x})$  has infinite support<sup>6</sup>. The form (6.4.12) may be convenient for computation in some cases.

Although the area density (6.4.11) is computed easily, it does not appear to be based on any optimality criterion; several methods based on such criteria have been found to yield estimates  $\hat{f}$  with reduced RMS error, so we focus on such methods next.

e,four,dens,con

<sup>&</sup>lt;sup>6</sup> The Kaiser Bessel windows that are used for  $C(\vec{\nu})$  in gridding are close approximations to the **prolate spheroidal wave functions** [121, 122]. The corresponding functions  $c(\vec{x})$  have minimal energy outside of  $\mathcal{D}$  so the approximation (6.4.12) should be reasonable.

#### s, four, dens, im 6.4.2.3 Optimality criteria: image domain (s, four, dens, im)

Several "optimal" methods for choosing the sampling density compensation vector w have been proposed, most of which are based on analyzing a discrete-discrete version of the estimator (6.4.2), *i.e.*, corresponding to the model (6.2.3). Most analyses consider the case of noiseless data, *i.e.*, when  $y = \bar{y} \triangleq \mathsf{E}[y] = \mathcal{A}f$ .

We begin by considering image-domain optimality criteria. For accurate reconstruction from noiseless data using (6.4.3), we would like to choose w such that [123, eqn. 13]

$$f \approx \mathsf{E}\Big[\hat{f}\Big] = \mathcal{Z} D(w) \mathsf{E}[y] = \mathcal{Z} D(w) \mathcal{A}f.$$
(6.4.13)

**6.4.2.3.1** Object-dependent image-domain DCF Although f is unknown in practice, for the purposes of establishing an ultimate performance bound in simulations, we could choose w by the following optimization criterion:

$$\hat{\boldsymbol{w}} = \underset{\boldsymbol{w}}{\operatorname{arg\,min}} \left\| f - \boldsymbol{\mathcal{Z}} \boldsymbol{D}(\boldsymbol{w}) \boldsymbol{\mathcal{A}} f \right\|_{\boldsymbol{\mathcal{K}}^{-1/2}}^{2}, \qquad (6.4.14)$$

where  $\mathcal{K}$  is a user-selected, positive-definite weighting *operator* on  $\mathcal{L}_2(\mathcal{D})$ . This  $\hat{w}$  is the ultimate **object-dependent**, **image-domain** choice, perhaps useful as a benchmark for comparing other methods.

Simplifying the cost function (6.4.14) yields

$$\begin{split} \|f - \mathcal{Z} \boldsymbol{D}(\boldsymbol{w}) \, \bar{\boldsymbol{y}}\|_{\mathcal{K}^{-1/2}}^2 &= \|f\|_{\mathcal{K}^{-1/2}}^2 - 2 \operatorname{real} \left\{ \langle \mathcal{K}^{-1} \, f, \, \mathcal{Z} \, \boldsymbol{D}(\boldsymbol{w}) \, \bar{\boldsymbol{y}} \rangle \right\} + \|\mathcal{Z} \, \boldsymbol{D}(\boldsymbol{w}) \, \bar{\boldsymbol{y}}\|_{\mathcal{K}^{-1/2}}^2 \\ &= \|f\|_{\mathcal{K}^{-1/2}}^2 - 2 \operatorname{real} \left\{ \langle \mathcal{Z}^* \mathcal{K}^{-1} \, f, \, \boldsymbol{D}(\boldsymbol{w}) \, \bar{\boldsymbol{y}} \rangle \right\} + \langle \mathcal{K}^{-1} \mathcal{Z} \, \boldsymbol{D}(\boldsymbol{w}) \, \bar{\boldsymbol{y}}, \, \mathcal{Z} \, \boldsymbol{D}(\boldsymbol{w}) \, \bar{\boldsymbol{y}} \rangle \\ &= \|f\|_{\mathcal{K}^{-1/2}}^2 - 2 \operatorname{real} \left\{ w' \operatorname{diag} \{ \bar{y}_i \}' \, \mathcal{Z}^* \mathcal{K}^{-1} \, f \right\} + w' \operatorname{diag} \{ \bar{y}_i \}' \, \mathcal{J} \operatorname{diag} \{ \bar{y}_i \} \, w, \end{split}$$

where J was defined in (6.4.6). So the gradient with respect to w is

2 diag
$$\{\bar{y}_i\}' \left( \boldsymbol{J} \operatorname{diag}\{\bar{y}_i\} \, \boldsymbol{w} - \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}^{-1} \, f 
ight)$$
 .

The (unconstrained, possibly complex) minimizer satisfies

$$J\operatorname{diag}\{\mathcal{A}f\}w = \mathcal{Z}^*\mathcal{K}^{-1}f.$$
(6.4.15)

If we choose  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ , then  $J = \tilde{K}$ , and the above system of equations always has a solution since  $\tilde{K}$  is invertible for distinct frequency samples. One can solve easily an equation of the form  $\tilde{K} u = v$  using the PCG iteration with a gridding approximation. If J is nearly singular, or if the noiseless data  $\bar{y}$  has zeros, then there are multiple density compensation vectors w that will minimize (6.4.14) nearly equally well. In that case, one could use the flexibility afforded by having multiple solutions to enforce other desirable constraints, such as nonnegativity:  $w \succeq 0$ , or some criterion related to the noise properties, as described in §6.4.2.1.

Rearranging (6.4.15), if J is invertible then the minimizer satisfies  $D(w) \mathcal{A} f = J^{-1} \mathcal{Z}^* \mathcal{K}^{-1} f$ , so the minimum (weighted) error in (6.4.14) is

$$\begin{split} \left\| f - \mathcal{Z} J^{-1} \mathcal{Z}^* \mathcal{K}^{-1} f \right\|_{\mathcal{K}^{-1/2}} &= \left\| \mathcal{K}^{1/2} \left( \mathcal{I} - \mathcal{K}^{-1/2} \mathcal{Z} J^{-1} \mathcal{Z}^* \mathcal{K}^{-1/2} \right) \mathcal{K}^{-1/2} f \right\|_{\mathcal{K}^{-1/2}} \\ &= \left\| \mathcal{P}_{\mathcal{K}^{-1/2} \mathcal{Z}}^{\perp} \mathcal{K}^{-1/2} f \right\|. \end{split}$$

When  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ , this is the same (weighted) error produced by the MNLS method (6.3.2) for noiseless data and when  $f_0 = 0$ .

**6.4.2.3.2** Min-max image-domain DCF Because f is unknown in practice, it is necessary to design the DCF without considering a specific object f. One possibility would be to minimize the worst-case error of the approximation (6.4.13) over some set  $\mathcal{F}$  of possible objects by the following min-max optimality criterion:

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \max_{f \in \mathcal{F}} \| f - \mathcal{Z} \boldsymbol{D}(\boldsymbol{w}) \mathcal{A} f \|_{\mathcal{K}_{2}^{-1/2}}, \qquad (6.4.16)$$

where  $\mathcal{K}_2$  denotes a user-selected, positive-definite image domain weighting operator on  $\mathcal{L}_2(\mathcal{D})$ . If one were to choose the set

$$\mathcal{F} = \left\{ f \in \mathcal{L}_2(\mathcal{D}) : \left\| f \right\|_{\mathcal{K}_1^{-1/2}} \le 1 \right\},\tag{6.4.17}$$

e,four,dens,im,w,equality

e,four,f<=]

then the min-max criterion (6.4.16) would be equivalent to the following optimization criterion:

$$\hat{\boldsymbol{w}} = \operatorname*{arg\,min}_{\boldsymbol{w}} \| \mathcal{K}_2^{-1/2} \left( \mathcal{I} - \mathcal{Z} \boldsymbol{D}(\boldsymbol{w}) \, \mathcal{A} \right) \mathcal{K}_1^{1/2} \| .$$
(6.4.18)

However, because the null space of  $\mathcal{AK}_1^{1/2}$  is nonempty, one can show that  $\|\mathcal{K}_2^{-1/2}(\mathcal{I} - \mathcal{Z}D(w)\mathcal{A})\mathcal{K}_1^{1/2}\|$  is independent of w. So one must choose a set  $\mathcal{F}$  that is orthogonal to the null space of  $\mathcal{A}$ . One possibility is to choose  $\mathcal{F} = \mathcal{R}_{\mathcal{Z}} \cap \left\{ f \in \mathcal{L}_2(\mathcal{D}) : \|f\|_{\mathcal{K}_1^{-1/2}} \leq 1 \right\}$ , in which case the min-max criterion (6.4.16) simplifies (Problem 6.16) to

$$\underset{\boldsymbol{w}}{\operatorname{arg\,min}} \| \boldsymbol{J}_{2}^{1/2} \left( \boldsymbol{I} - \boldsymbol{D}(\boldsymbol{w}) \, \boldsymbol{\mathcal{AZ}} \right) \boldsymbol{J}_{1}^{-1/2} \| , \qquad (6.4.19)$$

where, cf. (6.4.6), we define

s,four,dens,im,frob,

$$\boldsymbol{J}_1 \triangleq \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}_1^{-1} \boldsymbol{\mathcal{Z}}, \quad \boldsymbol{J}_2 \triangleq \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}_2^{-1} \boldsymbol{\mathcal{Z}}.$$

Computing this min-max solution efficiently appears to be a challenging open problem.

An alternative condition was studied by Choi and Munson [124] in the context of band-limited signal interpolation. Because components of f in the null space of  $\mathcal{A}$  are unrecoverable by linear methods, we should restrict attention to objects f that are not in that null space, specifically:  $\mathcal{F} = \{f \in \mathcal{L}_2(\mathcal{D}) \cap \mathcal{N}_{\mathcal{A}}^{\perp} : ||f|| \leq 1\}$ . If we assume that  $\mathcal{Z} = \mathcal{A}^*$ , then the solution given in [124, eqn. (31)] applies, which requires that one choose w to cluster the eigenvalues of  $\tilde{K} D(w)$  near unity, an impractical procedure for large  $n_d$ . Finding such a solution for more general  $\mathcal{Z}$  appears to be an open problem. (See (6.4.41) too.)

**6.4.2.3.3 Weighted Frobenius norm criterion in data domain** A simpler alternative to (6.4.19) would be to use a (possibly weighted) **Frobenius norm** criterion:

$$\underset{\boldsymbol{w}}{\operatorname{arg\,min}} \| \boldsymbol{W}_{1}^{1/2} \left( \boldsymbol{I} - \boldsymbol{D}(\boldsymbol{w}) \, \boldsymbol{\mathcal{AZ}} \right) \boldsymbol{W}_{2}^{1/2} \|_{\operatorname{Frob}}, \tag{6.4.20}$$

where  $W_1$  and  $W_2$  are user-selected, positive-semidefinite  $n_d \times n_d$  weighting matrices. By Problem 6.3, the minimizer satisfies

$$\begin{aligned} \boldsymbol{H} \hat{\boldsymbol{w}} &= \boldsymbol{v} \\ \boldsymbol{H}_{li} &= [\boldsymbol{W}_1]_{li} [\boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{Z}} \boldsymbol{W}_2 \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{A}}^*]_{il} \\ \boldsymbol{v}_i &= [\boldsymbol{W}_1 \boldsymbol{W}_2 \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{A}}^*]_{ii}, \qquad i = 1, \dots, n_d. \end{aligned}$$

In particular, if we choose  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ , then the solution simplifies to

$$\begin{aligned} H_{li} &= [\boldsymbol{W}_1]_{li} [\boldsymbol{K} \boldsymbol{W}_2 \boldsymbol{K}]_{il} \\ v_i &= [\boldsymbol{W}_1 \boldsymbol{W}_2 \tilde{\boldsymbol{K}}]_{ii}, \qquad i = 1, \dots, n_d. \end{aligned}$$

$$\overset{\text{e, four, dens, im, dat, frob,}}{\textbf{(6.4.22)}}$$

There may be a variety of choices of frequency domain weighting matrices  $W_1$  and  $W_2$  that lead to useful methods. (See Problem 6.5.) Instead of pursuing these, §6.4.2.3.4 considers another optimality formulation that turns out to be somewhat more general.

**6.4.2.3.4 Weighted Frobenius norm criterion in image domain** Another alternative to (6.4.18) is the following object-domain weighted Frobenius norm criterion:

$$\hat{\boldsymbol{w}} = \operatorname*{arg\,min}_{\boldsymbol{w}} \| \boldsymbol{\mathcal{W}}_{1}^{1/2} \left[ \boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{Z}} \boldsymbol{D}(\boldsymbol{w}) \boldsymbol{\mathcal{A}} \right] \boldsymbol{\mathcal{W}}_{2}^{1/2} \|_{\mathrm{Frob}}, \tag{6.4.23}$$

where  $W_1$  and  $W_2$  are user-selected, positive-semidefinite object-domain weighting operators. Such weighting could be useful when different image regions (*e.g.*, near borders) have different importance; see [16] for Schwab's weighting function. The unweighted version of this Frobenius norm was used implicitly to choose  $\hat{w}$  via iterative minimization methods in [94]. However, one can find the following explicit closed-form solution by Problem 6.3 (*cf.* [125, p. 311]):

$$\begin{aligned} H\hat{w} &= v \\ H_{li} &= [\mathcal{Z}^* \mathcal{W}_1 \mathcal{Z}]_{li} [\mathcal{A} \mathcal{W}_2 \mathcal{A}^*]_{il} \\ v_i &= [\mathcal{Z}^* \mathcal{W}_1 \mathcal{W}_2 \mathcal{A}^*]_{ii}, \qquad i = 1, \dots, n_d. \end{aligned}$$
(6.4.24)

#### © J. Fessler. [license] December 10, 2018

Although this solution is explicit, it would still appear to require considerable computation in general.

If we choose  $\mathcal{W}_1 = \mathcal{K}_1^{-1} \mathcal{Z} J_1^{-1} W_1 J_1^{-1} \mathcal{Z}^* \mathcal{K}_1^{-1}$  and  $\mathcal{W}_2 = \mathcal{Z} W_2 \mathcal{Z}^*$ , then the solution (6.4.24) becomes identical to (6.4.21). So the frequency domain Frobenius criterion (6.4.20) is a special case of the object-domain Frobenius criterion (6.4.23). One can show that the solutions (6.4.24) are in fact more general than (6.4.21) (see Problem 6.18), but whether that generality leads to useful methods is an open problem. We now consider several special cases for the choices of  $\mathcal{W}_1, \mathcal{W}_2$ , and  $\mathcal{Z}$ .

• For the choice  $\mathcal{W}_1 = \mathcal{A}^* \tilde{\mathcal{K}}^{-2} \mathcal{A}, \mathcal{W}_2 = \mathcal{K} \mathcal{A}^* \tilde{\mathcal{K}}^{-2} \mathcal{A} \mathcal{K}, \text{ and } \mathcal{Z} = \mathcal{K} \mathcal{A}^*, \text{ we find } H = I \text{ and } v_i = \left| \tilde{\mathcal{K}}^{-1} \right|_{i:}, \text{ so}$ the solution is the diagonal of the inverse of the weighted crosstalk matrix:

$$\hat{w}_i = \left[\tilde{\boldsymbol{K}}^{-1}\right]_{ii}$$

This matrix inverse seems impractical.

• If we choose  $\mathcal{W}_1 = \mathcal{K}^{-1}$ ,  $\mathcal{W}_2 = \mathcal{K}\mathcal{A}^*\tilde{\mathcal{K}}^{-2}\mathcal{A}\mathcal{K}$ , and  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ , then  $H = \mathsf{diag}\left\{\tilde{K}_{ii}\right\}$  and v = 1 so the solution is the reciprocal of the diagonal of the weighted crosstalk matrix:

$$\hat{w}_i = 1/K_{ii}$$

If we choose  $\mathcal{K} = \mathcal{D}(c(\vec{x}))$ , then  $\tilde{K}_{ii}$  is the constant value  $(C * S)(\vec{0})$ , so this choice is useless. • For the choice  $\mathcal{W}_1 = \mathcal{K}^{-1}$ ,  $\mathcal{W}_2 = \mathcal{K}\mathcal{A}^*\tilde{K}^{-1}\mathbf{1}\mathbf{1}'\tilde{K}^{-1}\mathcal{A}\mathcal{K}$ , and  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ , then  $H = \tilde{K}$  and v = 1 so the minimizer is

$$\hat{w} = K^{-1} \mathbf{1}$$

Although this expression appears to involve a matrix inverse, it can be computed simply by solving the system of equations e,four,dens,im,Kw=

$$\tilde{K}\hat{w} = 1 \tag{6.4.25}$$

using an iterative algorithm like the conjugate gradient (CG) method, so it can be made practical. In particular, if we make the choice

$$\mathcal{K} = \mathcal{D}(|c(\vec{x})|^2),$$

where  $c(\vec{x}) s(\vec{x}) \approx c(\vec{x})$ , then by using (6.1.17) we see that

s,four,dens,pint

$$\tilde{K}_{il} \approx (C * C) (\vec{\nu}_i - \vec{\nu}_l).$$

Because this is a (block) banded matrix, one can solve (6.4.25) iteratively fairly easily.

• If we choose  $\mathcal{W}_1 = \mathcal{K}^{-1}$ ,  $\mathcal{W}_2 = \mathcal{K}$ , and  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ , then  $H_{li} = \left|\tilde{K}_{li}\right|^2$  and  $v_i = \tilde{K}_{ii}$ . In particular, if we choose  $\mathcal{K} = \mathcal{D}(|c(\vec{x})|^2)$ , then  $H_{il} \approx |(C * C)(\vec{\nu_i} - \vec{\nu_l})|^2$  and  $v_i \approx (C * C)(0)$ . This too is a (block) banded system of

equations; developing an efficient implementation of this solution is an open problem. Different choices for  $\mathcal{W}_1$ ,  $\mathcal{W}_2$ , and  $\mathcal{Z}$  lead to quite different choices for  $\hat{w}$ . These could be compared for a given object of interest f using (6.4.14).

6.4.2.3.5 Pseudo-inverse criteria (s,four,dens,piny) If f were in the intersection of the range of  $\mathcal{A}^*$  and the range of  $\boldsymbol{Z}$ , then the approximation (6.4.13) would be exact if we could choose  $\boldsymbol{D}$  to be a  $n_{\rm d} \times n_{\rm d}$  matrix of the following imposing form:

$$oldsymbol{D}_0 = oldsymbol{W}_1oldsymbol{\mathcal{Z}}^* [oldsymbol{\mathcal{Z}} oldsymbol{W}_1oldsymbol{\mathcal{Z}}^*]^\dagger [oldsymbol{\mathcal{A}}^* oldsymbol{W}_2oldsymbol{\mathcal{A}}]^\dagger oldsymbol{\mathcal{A}}^* oldsymbol{W}_2,$$

where  $W_1, W_2 \in \mathbb{C}^{n_d \times n_d}$  are user-selected, positive-definite, Hermitian weighting matrices. A similar expression was suggested in [125, eqn. (6)]. This  $D_0$  is not diagonal, so one could attempt to find a diagonal approximation to it by an optimality criterion such as the following:

$$\underset{\boldsymbol{w}}{\operatorname{arg\,min}} \| \boldsymbol{D}_0 - \boldsymbol{D}(\boldsymbol{w}) \| _{\boldsymbol{W}_3^{1/2}},$$

where  $W_3 \in \mathbb{C}^{n_d \times n_d}$  is another user-selected, positive-definite, Hermitian weighting matrix. It is somewhat difficult to see what the most suitable choices for  $W_1$ ,  $W_2$ , and  $W_3$  would be, and difficult to solve the minimization problem because of the pseudo-inverse.

The problem becomes more tractable if instead we use a Frobenius norm:

$$\hat{\boldsymbol{w}} = \operatorname*{arg\,min}_{\boldsymbol{w}} \| \boldsymbol{W}_3^{1/2} (\boldsymbol{D}(\boldsymbol{w}) - \boldsymbol{D}_0) \boldsymbol{W}_4^{1/2} \|_{\mathrm{Frob.}}$$

Problem 6.3 shows that the minimizer satisfies

$$\begin{split} H\hat{w} &= v \\ H_{li} &= [W_3]_{li} [W_4]_{il} \\ v_i &= [W_3 D_0 W_4]_{ii}, \quad i = 1, \dots, n_d. \end{split} \tag{6.4.26}$$

If either  $W_3$  or  $W_4$  are diagonal, then so is H, and the solution simplifies to

$$\hat{w}_i = rac{[W_3 D_0 W_4]_{ii}}{[W_3]_{ii} [W_4]_{ii}}.$$

For  $W_3 = W_4 = I$ , this is simply the diagonal of  $D_0$ . In particular, for the case  $\mathcal{Z} = \mathcal{A}^*$ , using the approximation (6.1.15) yields

$$[\boldsymbol{D}_0]_{ii} = \boldsymbol{e}'_i \boldsymbol{\mathcal{A}} [\boldsymbol{\mathcal{A}}^* \boldsymbol{\mathcal{A}}]^{\dagger} [\boldsymbol{\mathcal{A}}^* \boldsymbol{\mathcal{A}}]^{\dagger} \boldsymbol{\mathcal{A}}^* \boldsymbol{e}_i \approx \boldsymbol{e}'_i \boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{F}}^{-1} \boldsymbol{\mathcal{D}} \left( \left| H(\vec{\nu}) \right|^2 \right)^{\dagger} \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{A}}^* \boldsymbol{e}_i = \int \frac{|S(\vec{\nu} - \vec{\nu}_i)|^2}{|H(\vec{\nu})|^2} \, \mathrm{d}\vec{\nu}_i$$

where  $H(\vec{\nu})$  was defined in (6.1.14). Evaluating this choice, *e.g.*, by using NUFFTs, is an open problem.

Sedarat and Nishimura proposed a closely related optimality criterion [125] (in the context of gridding), that we generalize here. Let  $\hat{f}_{ref} = \mathcal{R} y$  denote a good "reference" reconstruction method, for example, we could use  $\mathcal{R} = \mathcal{A}^* D_0$ . The difference between the reference reconstruction and the conjugate-phase estimate based on a diagonal D(w) is  $\hat{f} - \hat{f}_{ref} = \mathcal{Z} D(w) y - \mathcal{R} y = (\mathcal{Z} D(w) - \mathcal{R})y$ . So a natural criterion for w is to minimize the following weighted Frobenius norm:

$$\| \mathcal{K}_{1}^{1/2} (\mathcal{Z} D(w) - \mathcal{R}) W_{3}^{1/2} \|_{\text{Frob}} = \| \mathcal{K}_{1}^{1/2} \mathcal{R} W_{3}^{1/2} - \mathcal{K}_{1}^{1/2} \mathcal{Z} D(w) W_{3}^{1/2} \|_{\text{Frob}}, \qquad (6.4.27)$$

where  $\mathcal{K}_1^{1/2}$  and  $W_3^{1/2}$  are any Hermitian symmetric, positive definite image domain operator and  $n_d \times n_d$  matrix, respectively. Problem 6.3, which generalizes [125, eqn. (8)], shows that the minimizer satisfies

$$\begin{split} \boldsymbol{H} \hat{\boldsymbol{w}} &= \boldsymbol{v} \\ \boldsymbol{H}_{li} &= [\boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}_1 \boldsymbol{\mathcal{Z}}]_{li} [\boldsymbol{W}_3]_{il} \\ \boldsymbol{v}_i &= [\boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}_1 \boldsymbol{\mathcal{R}} \boldsymbol{W}_3]_{ii}, \quad i = 1, \dots, n_d. \end{split}$$

$$\begin{aligned} &\stackrel{\text{e, four, dens, im, frob, wtd, s}}{(6.4.28)} \end{split}$$

In particular, if  $W_3$  is any diagonal matrix, then H is also diagonal and

$$\hat{w}_i = \frac{[\boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}_1 \boldsymbol{\mathcal{R}}]_{ii}}{[\boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{K}}_1 \boldsymbol{\mathcal{Z}}]_{ii}}$$

Interestingly, for  $\mathcal{R} = \mathcal{K}_2^{1/2} \left[ A \mathcal{K}_2^{1/2} \right]^{\dagger}$  and  $W_3 = \mathcal{A} \mathcal{K}_2 \mathcal{A}^*$ , the solution (6.4.28) simplifies to (6.4.24). Apparently (6.4.27) is more general than (6.4.23). However, the obvious choices for  $W_3$  and  $\mathcal{R}$  lead to solutions shown in §6.4.2.3.4, so whether this additional generality is useful is an *open problem*.

#### s, four, dens, psf 6.4.2.4 Optimality criteria: PSF (s, four, dens, psf)

If  $f(\vec{x}) = \delta(\vec{x})$ , a Dirac impulse at the center of the spatial coordinates, then Af = 1, the vector of  $n_d$  ones, in which case

$$\mathcal{Z} D(w) \mathcal{A} f = \mathcal{Z} w,$$

which is sometimes called "the" point-spread function (PSF) of the reconstruction method (6.4.3). This terminology is imprecise because the method (6.4.3) is *not* exactly shift invariant in general, due to the  $s(\vec{x})$  term in (6.4.2). So in fact there is a somewhat different PSF for each spatial location. Nevertheless, for good spatial resolution, we would like this PSF to closely approximate a Dirac impulse. However, the Dirac impulse is not in  $\mathcal{L}_2(\mathcal{D})$ , so we must choose a square-integrable "target" PSF for optimizing w. (In some cases one can consider a Dirac impulse by limiting arguments [126].) A logical choice of target impulse response would be the function whose spectrum is unity over a set C that closely approximates the convex hull of the sample locations { $\vec{v}_i$ } and is zero otherwise. For example, for a spiral sampling pattern, C would be a disk. We could then define the target PSF as follows

$$h(\vec{\mathbf{x}}) = \int_{\mathcal{C}} e^{i2\pi\vec{\nu}\cdot\vec{\mathbf{x}}} \,\mathrm{d}\vec{\nu} \,.$$

It follows that  $Ah \approx 1$ ; the reason this is only *approximate* is that the target PSF h is not exactly support limited to D. (An alternative target PSF would be the function h that is support limited to D that is as close to unity over C and to zero over the complement of C as possible.)

Having chosen such a target PSF, we can then optimize w by minimizing the following WLS cost function:

$$\|h - \mathcal{Z}\boldsymbol{w}\|_{\boldsymbol{\mathcal{W}}_{0}^{1/2}}^{2} \equiv \boldsymbol{w}'\tilde{\boldsymbol{J}}\boldsymbol{w} - 2\operatorname{real}\{\boldsymbol{w}'\boldsymbol{u}\}, \tag{6.4.29}$$

for some user-selected weighting operator  $\mathcal{W}_0$ , where

$$\tilde{\boldsymbol{J}} \triangleq \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{W}}_0 \boldsymbol{\mathcal{Z}}$$
(6.4.30)

is a  $n_{\rm d} \times n_{\rm d}$  relative of the weighted Fourier crosstalk matrix  $\hat{K}$  in (6.3.3), and where

$$\boldsymbol{u} \triangleq \boldsymbol{\mathcal{Z}}^* \boldsymbol{\mathcal{W}}_0 h.$$

If  $\mathcal{Z} = \mathcal{A}^*$  and  $\mathcal{W}_0 = \mathcal{D}(|c(\vec{x})|^2)$ , then  $\tilde{J}$  is a symmetric banded matrix with entries

$$\tilde{J}_{il} = (C * C)(\vec{\nu}_i - \vec{\nu}_l)$$

The gradient (see \$28.2) of this cost function with respect to w is

$$ilde{J}w - u$$

For unconstrained minimization of (6.4.29), the optimal w satisfies

$$\hat{J}\hat{w} = u, \tag{6.4.31}$$

*i.e.*, is given by

$$\hat{oldsymbol{w}} = ilde{oldsymbol{J}}^{-1} oldsymbol{u} = \left[oldsymbol{\mathcal{Z}}^* oldsymbol{\mathcal{W}}_0 oldsymbol{\mathcal{Z}}
ight]^{-1} oldsymbol{\mathcal{Z}}^* oldsymbol{\mathcal{W}}_0 h$$

Because  $\tilde{J}$  is too large to invert directly, iterative methods are needed.

If  $\mathcal{Z} = \mathcal{W}_0^{-1} \mathcal{A}^*$  and we choose  $\mathcal{W}_0 = \mathcal{K}^{-1}$ , then  $u = \mathcal{Z}^* \mathcal{W}_0 h = \mathcal{A} h \approx 1$ , and  $\mathcal{Z}^* \mathcal{W}_0 \mathcal{Z} = \mathcal{A} \mathcal{K}^{-1} \mathcal{A}^* = \tilde{K}$ , so the desired w corresponds to "solving" the following system of equations

$$\tilde{K}\hat{w} = 1. \tag{6.4.32}$$

Several methods related to solving (6.4.32) or an approximation thereof have been proposed in the literature. However, if  $\mathcal{Z}$  is a crude approximation to  $\mathcal{A}^*$ , then  $u \neq 1$ . Most published methods have ignored this distinction, the importance of which is an open problem.

**6.4.2.4.1** Pipe and Menon's iteration In the context of gridding, Pipe and Menon [17] observed (for u = 1) that dividing both sides of (6.4.31) into the vector w yields the following equality

$$\hat{oldsymbol{w}}=\hat{oldsymbol{w}}\odot\left[oldsymbol{u}\oslash\left( ilde{oldsymbol{J}}\hat{oldsymbol{w}}
ight)
ight],$$

where  $\oslash$  and  $\odot$  denote element-wise division and multiplication (Hadamard product). This equality suggests the following multiplicative fixed-point iteration

$$\boldsymbol{w}^{(n+1)} = \boldsymbol{w}^{(n)} \odot \left[ \boldsymbol{u} \oslash \left( \tilde{\boldsymbol{J}} \boldsymbol{w}^{(n)} \right) \right], \qquad (6.4.33)$$

I am unaware of any proof of convergence for this iteration, nor does this iteration appear to be related to any specific optimality criterion. It does have the property that if the initial vector  $w^{(0)}$  is nonnegative, then all subsequent iterations  $w^{(n)}$  will be nonnegative if the elements of  $\tilde{J}$  and u are nonnegative and real.

e,four,dens,Jh

e,four,dens,Jw=u

e,four,dens,Kw=]

e,four,dens,pipe

**6.4.2.4.2** An EM iteration The form of (6.4.33) is suggestively close to the EM algorithm for emission tomography described in Chapter 16. In the notation used here, that EM iteration is:

$$\boldsymbol{w}^{(n+1)} = \boldsymbol{w}^{(n)} \odot \left[ \tilde{\boldsymbol{J}} \left( \boldsymbol{u} \oslash \left[ \tilde{\boldsymbol{J}} \boldsymbol{w}^{(n)} \right] \right) \right] \oslash \left( \tilde{\boldsymbol{J}} \boldsymbol{1} \right).$$
(6.4.34)

Normally this type of iteration is associated with Poisson measurement noise, but it also has been proposed for problems with nonnegativity constraints and noiseless "data" [127, 128]. This iteration *is* associated with an optimality criterion:

$$\min_{\boldsymbol{w} \succeq \boldsymbol{0}} \sum_{i=1}^{n_{\mathrm{d}}} \kappa \Big( 1, \Big[ \tilde{\boldsymbol{J}} \boldsymbol{w} \Big]_i \Big),$$

where  $\kappa$  denotes the Kullback-Leibler divergence (16.4.11). However, the algorithm (6.4.34) requires roughly twice as much computation per iteration as (6.4.33). If there is a nonnegative  $\hat{w}$  that satisfies (6.4.32), then (6.4.34) will converge to  $\hat{w}$  [129], whereas the most that we can state about (6.4.33) is that  $\hat{w}$  is a fixed point.

**6.4.2.4.3** Qian et al.'s iteration Considering (6.4.32), Qian et al. [130] proposed (for u = 1) the following additive fixed-point iteration:

$$\boldsymbol{w}^{(n+1)} = \boldsymbol{w}^{(n)} + \alpha \left( \boldsymbol{u} - \tilde{\boldsymbol{K}} \boldsymbol{w}^{(n)} \right)$$

with the step size  $\alpha$  chosen empirically. Although it was not derived explicitly as such, we see that this iteration is gradient descent of (6.4.29) with  $\mathcal{Z} = \mathcal{W}_0^{-1} \mathcal{A}^*$ . For a square FOV,  $\tilde{K}$  consists of samples of sinc functions, and Qian et al. found that truncating the sinc (to reduce computation) degraded  $\hat{w}$ . They also had to make certain *ad hoc* modifications near the densely sampled origin of a spiral sampling pattern. They reported modest improvements relative to the method of Pipe and Menon [17] in (6.4.33). Apparently their algorithm did not enforce the constraint  $w \succeq 0$ .

Dwork et al. [131] also proposed using LSQR to solve the same (unweighted) least-squares criterion (6.4.29), again with no mention of a nonnegativity constraint.

**6.4.2.4.4** Samsonov et al.'s iteration Samsonov et al. [132] proposed (for u = 1) a variation of this method that enforced the constraint  $w \succeq 0$  using a projected preconditioned steepest descent algorithm (*cf.* §11.3). Essentially, that algorithm tries to solve the following optimization problem

$$\min_{\boldsymbol{w} \succeq \boldsymbol{0}} \left\| \tilde{\boldsymbol{J}} \boldsymbol{w} - \boldsymbol{u} \right\|_{\boldsymbol{W}^{1/2}}, \qquad (6.4.35)^{\text{e},\text{i}}$$

for the unweighted case W = I. They initialized the iteration using (6.4.11), and they found that the RMS error between  $\hat{f}$  and f was smaller when using (6.4.35) than when they used the method of Pipe and Menon [17] in (6.4.33).

**6.4.2.4.5** Alternative PSF-based methods None of the four preceding algorithms are exactly equivalent to solving the original minimization problem (6.4.29) subject to a nonnegativity constraint. Instead, they all aim to approximate the *unconstrained* solution (6.4.32). A more principled approach is the following:

$$\hat{\boldsymbol{w}} = \underset{\boldsymbol{w} \in \mathbb{R}^{n_{d}}, \ \boldsymbol{w} \succeq \boldsymbol{0}}{\arg\min} \Psi(\boldsymbol{w}), \qquad \Psi(\boldsymbol{w}) = \frac{1}{2} \boldsymbol{w}' \tilde{\boldsymbol{J}} \boldsymbol{w} - \boldsymbol{w}' \operatorname{real}\{\boldsymbol{u}\} + \beta \frac{1}{2} \|\boldsymbol{w}\|^{2}, \qquad (6.4.36)^{\text{effect}}$$

where the regularizer  $\beta \|w\|^2$  is desirable when  $\tilde{J}$  is poorly conditioned. This regularizer is motivated by the noise considerations in §6.4.2.1. Finding a good iterative algorithm to minimize this cost function subject to the nonnegativity constraint, for example, by using the optimization methods described in Chapter 14, is an interesting *open problem*. A natural approach is **FISTA** because of the nonnegativity constraint. Given the availability of approximate solutions, it should be possible to find a good preconditioner.

The criterion (6.4.29) was developed only for the PSF located at the center. An interesting open problem is to formulate criteria that involve all possible PSF locations, *e.g.*, using a min-max approach or averaging to combine.

The most common choice for the reconstructor  $\mathcal{Z}$  is the gridding method of §6.4.3.5; finding iterative methods for computing w in that context specifically is an interesting open problem.

Yet another approach is to minimize the maximum PSF sidelobe level while constraining mainlobe width, using linear matrix inequality constraints. [133]. Comparing this approach to the many alternatives described here is an open problem.

#### s, four, dens, freq 6.4.2.5 Optimality criteria: frequency domain (s, four, dens, freq)

Since the measurements are recorded in the frequency domain, it seems desirable for the spectrum of the estimate f to approximate closely the frequency samples, at least in the absence of noise. In other words, we would like

$$\mathcal{A}f = \mathcal{AZ}D(w)y \approx y$$

or equivalently

$$\boldsymbol{M}\boldsymbol{D}(\boldsymbol{w})\boldsymbol{y} = \boldsymbol{M}\operatorname{diag}\{y_i\}\boldsymbol{w} \approx \boldsymbol{y},\tag{6.4.37}$$

where we define  $M = \mathcal{A}\mathcal{Z}$ . This criterion is a kind of "self consistency."

The desired equality (6.4.37) suggests the following data-dependent (and hence impractical) optimality criterion:

$$\min_{\mathbf{W}} \| \boldsymbol{y} - \boldsymbol{M} \operatorname{diag}\{y_i\} \, \boldsymbol{w} \|_{\boldsymbol{W}^{1/2}} \,. \tag{6.4.38}$$

If M is invertible, which certainly is the case for distinct frequencies with  $\mathcal{Z} = \mathcal{A}^*$ , then the "optimal" data-dependent density compensation vector w is

$$\hat{\boldsymbol{w}} = \operatorname{diag}\{1/y_i\} \boldsymbol{M}^{-1} \boldsymbol{y},$$

assuming each  $y_i \neq 0$ . If  $y_i = 0$  for any *i*, then there will be multiple optimal *w*, and we could replace diag $\{1/y_i\}$  with a pseudo-inverse. Alternatively, we could introduce regularization as in (6.4.36), leading to the (unconstrained) solution

$$\hat{oldsymbol{w}} = \left[ \mathsf{diag}\{y_i\}' oldsymbol{M}' oldsymbol{W} oldsymbol{M} \, \mathsf{diag}\{y_i\} + eta oldsymbol{I} 
ight]^{-1} \, \mathsf{diag}\{y_i\}' \, oldsymbol{M}' oldsymbol{W} oldsymbol{y}.$$

Recomputing w for each y using (6.4.38) would be inconvenient in practice. Is there a particular y that would be useful for design? A reasonable choice would be y = 1. This choice is equivalent to requiring spectrum of  $\hat{f}$  be consistent with the data in the case of constant measurements. This requirement seems very reasonable. Such a density compensation vector w must satisfy the equality (6.4.32) discussed previously from a PSF perspective. So that equality and the corresponding criterion (6.4.35) have an additional justification.

If we restrict attention to nonnegative density compensation vectors, then in general the desired equality (6.4.37) will not hold exactly for *any*  $w \succeq 0$ . Furthermore, this consistency criterion may be unreasonably restrictive for noisy measurements, and may even be undesirable, particularly in densely sampled regions of frequency space.

If we wanted to find a single coefficient vector w that provides consistency for any set of measurements y, then we must have

$$\boldsymbol{M}\,\boldsymbol{D}(\boldsymbol{w}) = \boldsymbol{I}.\tag{6.4.39}$$

In general, no such vector w exists, so we can only have approximate consistency, no matter how fancy an iterative algorithm is used to find w. Again a min-max optimality criterion is a natural extension of (6.4.38):

$$\min_{\boldsymbol{w}} \max_{\|\boldsymbol{y}\| \le 1} \|\boldsymbol{y} - \boldsymbol{M} \operatorname{diag}\{y_i\} \boldsymbol{w}\| = \min_{\boldsymbol{w}} \max_{\|\boldsymbol{y}\| \le 1} \|\boldsymbol{y} - \boldsymbol{M} \boldsymbol{D}(\boldsymbol{w}) \boldsymbol{y}\| = \min_{\boldsymbol{w}} \|\boldsymbol{I} - \boldsymbol{M} \boldsymbol{D}(\boldsymbol{w})\|_2.$$
(6.4.40)

This norm criterion appears inconvenient for routine computation.

An object-based alternative to the min-max criterion (6.4.40) is the following:

$$\min_{\boldsymbol{w}} \max_{\|f\| \leq 1} \|\boldsymbol{\mathcal{A}} f - \boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{Z}} \boldsymbol{D}(\boldsymbol{w}) \boldsymbol{\mathcal{A}} f\|_{\boldsymbol{\mathcal{W}}^{1/2}}.$$

This optimization problem appears to be even more complicated.

**6.4.2.5.1** Weighted Frobenius norm in data domain An alternative to (6.4.40) is the following weighted Frobenius norm motivated by (6.4.39), with regularization:

$$\Psi(\boldsymbol{w}) = \| \boldsymbol{W}_1^{1/2} (\boldsymbol{I} - \boldsymbol{M} \boldsymbol{D}(\boldsymbol{w})) \boldsymbol{W}_2^{1/2} \|_{\text{Frob}}^2 + \beta \| \boldsymbol{w} \|^2.$$
(6.4.41)

Generalizing Problem 6.3, the minimizer satisfies

---

$$egin{aligned} & [H+eta I]w = & v \ & H_{li} = & & [M'W_1M]_{li}[W_2]_{il} \ & v_i = & & [M'W_1W_2]_{ii}. \end{aligned}$$

e,four,dens,opt,

In particular, if  $W_2$  is a diagonal matrix, then

$$\hat{w}_{i} = \frac{[M'W_{1}]_{ii} [W_{2}]_{ii}}{[M'W_{1}M]_{ii} [W_{2}]_{ii} + \beta}.$$
(6.4.43)

This solution generalizes the Frobenius approach given in [124, eqn. (34)]. In general, neither this explicit solution nor iterative methods for the minimization appear to be practical for large  $n_{\rm d}$ .

Now consider the specific case where  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$  and  $W_2 = I$ . Then  $M = \tilde{K}$  and the solution simplifies to

$$\hat{w}_i = \frac{[\tilde{K}W_1]_{ii}}{[\tilde{K}W_1\tilde{K}]_{ii} + \beta}.$$
(6.4.44)

Choices for  $W_1$  like  $\tilde{K}^{-2}$ ,  $\tilde{K}^{-1}$ , and  $\tilde{K}^{-1}\mathbf{11}'\tilde{K}^{-1}$ , lead to the solutions presented in §6.4.2.3.4. If we choose  $W_1 = I$  and  $\mathcal{K} = \mathcal{D}(|c(\vec{x})|^2)$  then  $\tilde{K}_{il} = (C * C)(\vec{\nu}_i - \vec{\nu}_l)$ . Thus the solution h

f we choose 
$$W_1 = I$$
 and  $\mathcal{K} = \mathcal{D}(|c(\vec{x})|^2)$ , then  $K_{il} = (C * C)(\vec{\nu}_i - \vec{\nu}_l)$ . Thus the solution becomes

$$\hat{w}_{i} = \frac{\int |C(\vec{\nu})|^{2} \,\mathrm{d}\vec{\nu}}{\sum_{l} |(C * C)(\vec{\nu}_{i} - \vec{\nu}_{l})|^{2} + \beta} \approx \frac{\int |C(\vec{\nu})|^{2} \,\mathrm{d}\vec{\nu}}{\sum_{l} |C(\vec{\nu}_{i} - \vec{\nu}_{l})|^{2} + \beta}.$$

These solutions appear to be quite practical. Whether they work well is an *open problem*. A similar method akin to the choice  $c(\vec{x}) = s(\vec{x})$  was examined in [134]. In particular, in that case we have

$$\hat{w}_{i} = \frac{S\left(\vec{0}\right)}{\sum_{l=1}^{n_{\mathrm{d}}} \left|S\left(\vec{\nu}_{l} - \vec{\nu}_{i}\right)\right|^{2} + \beta}$$

Despite the infinite support of  $S(\vec{\nu})$ , one can evaluate the denominator summation efficiently using NUFFT operations [134], because

$$\sum_{l=1}^{n_{\rm d}} |S(\vec{\nu}_i - \vec{\nu}_l)|^2 = \int (s \star s)(\vec{\mathbf{x}}) \left[ \sum_{l=1}^{n_{\rm d}} e^{i2\pi \vec{\nu}_l \cdot \vec{\mathbf{x}}} \right] e^{-i2\pi \vec{\nu}_i \cdot \vec{\mathbf{x}}} \, \mathrm{d}\vec{\mathbf{x}} \, .$$

The inner bracketed expression is the adjoint of a NUFFT, and the outer integral, when discretized, is a NUFFT [50].

#### s, four, dens, summ 6.4.2.6 Density compensation summary (s, four, dens, summ)

There are many possible methods for choosing w for use with the non-iterative conjugate-phase reconstruction method (6.4.3). There are a few simple choices that can be computed easily, and there are many possible "optimal" choices.

We have described all of the methods in terms of the operator  $\mathcal{A}$  and in terms of a *fixed* reconstruction operator  $\mathcal{Z}$ . An interesting extension is to consider joint optimization of  $\mathcal{Z}$  and w, such as simultaneously optimizing the gridding kernel(s) and w [125].

Another intriguing extension would be to allow *space varying* sampling density compensation, *i.e.*, to replace (6.4.2) with the following estimator

$$\hat{f}(\vec{\mathbf{x}}) = \sum_{i=1}^{n_{\rm d}} w_i(\vec{\mathbf{x}}) y_i \, \mathrm{e}^{i 2\pi \vec{\nu}_i \cdot \vec{\mathbf{x}}} \,,$$

where now one must design the density compensation *functions*  $\{w_i(\vec{x})\}$ . Noll et al. [115] have proposed such a method in the context of MR imaging using spiral sampling patterns in the presence of magnetic field inhomogeneity. The extension to other sampling patterns is an interesting open problem; see [135, 136].

Another possibility would be to allow D to have a banded structure instead of being restricted to diagonal [124, 125]. Density compensation methods for under-sampled cases have also been investigated [137].

All of these possible extensions add further complications, and at some point one might as well abandon the conjugate-phase reconstruction method (6.4.3), and instead use iterative methods, such as the QPWLS approach described in §6.3.2, that conveniently do not require any density compensation!

For additional methods, see [138] [139] [140].

#### s, four, freq 6.4.3 Frequency-domain interpolation methods (s, four, freq)

The methods described above all attempt to estimate directly the object f from the data y. An alternative approach is to first use the measurements y to find an estimate  $\hat{F}(\vec{\nu})$  of the spectrum of f, and then estimate f by an inverse Fourier transform:

$$\hat{f}(\vec{\mathbf{x}}) = \int \hat{F}(\vec{\nu}) \, \mathrm{e}^{\imath 2 \pi \vec{\nu} \cdot \vec{\mathbf{x}}} \, \mathrm{d}\vec{\nu}$$

In practice, usually this integral is approximated by a finite sum. In particular, discretizing the integral using uniform spacing corresponds to an inverse FFT.

Many methods for estimating  $F(\vec{\nu})$  from y assume the data is noiseless, *i.e.*,  $y_i = F(\vec{\nu}_i)$ . Since f is assumed to be space limited to a bounded domain  $\mathcal{D}$ , estimating  $F(\vec{\nu})$  from a noiseless y is equivalent to the classical problem of interpolating band-limited signals from nonuniformly spaced samples, for which there is extensive literature *e.g.*, [59, 61, 109, 124, 126, 141–156].

The assumption that f is space limited is very realistic physically in many imaging applications. In contrast, the conventional assumption that signal samples are taken from a band-limited signal is often a mathematical convenience rather than an absolute physical constraint. Indeed, some interpolation methods include procedures for *estimating* the band width from the data [44].

Usually the estimate  $F(\vec{v})$  is a linear function of y, *i.e.*, it has the form

$$\hat{F}(\vec{\nu}) = \sum_{i=1}^{n_{\rm d}} R_i(\vec{\nu}) y_i$$
(6.4.45)

for some functions  $\{R_i(\vec{\nu})\}$ . In that case the corresponding image-domain estimate is "simply"

$$\hat{f}(\vec{\mathbf{x}}) = \sum_{i=1}^{n_{\rm d}} r_i(\vec{\mathbf{x}}) \, y_i$$

where  $r_i(\vec{x})$  is the inverse FT of  $R_i(\vec{\nu})$ . Various linear methods differ "only" in the choice of  $\{R_i(\vec{\nu})\}$ .

#### s, four, freq, block 6.4.3.1 Block spectral interpolation methods (s, four, freq, block)

In the general linear expression (6.4.45) above, for any frequency  $\vec{\nu}$  the estimate  $\hat{F}(\vec{\nu})$  could depend on all data values  $\{y_i\}$ . Such an approach would be practical only for problems with very few measurements (small  $n_d$ ). For large problems like those in imaging, usually we estimate  $F(\vec{\nu})$  for any given frequency  $\vec{\nu}$  using only the frequency samples  $\vec{\nu}_i$  in the local neighborhood of  $\vec{\nu}$ , *i.e.*, we replace (6.4.45) with a sum of the form

$$\hat{F}(\vec{\nu}) = \sum_{i \in \mathcal{I}_{\vec{\nu}}} R_i(\vec{\nu}) y_i, \tag{6.4.46}$$

where  $\mathcal{I}_{\vec{\nu}}$  is a *subset* of  $\{1, \ldots, n_d\}$ , such as the K neighbors in  $\{\vec{\nu}_i\}$  that are closest to  $\vec{\nu}$ , or perhaps the following

$$\mathcal{I}_{\vec{\nu}} = \{i \in \{1, \dots, n_{\mathrm{d}}\} : \|\vec{\nu} - \vec{\nu}_i\| \le \nu_{\mathrm{radius}}\}.$$

Such methods have been called **block interpolation** methods and proposals for their use date back at least to [157] in band-limited signal interpolation and to [4] in imaging.

For notational simplicity, we describe many of the interpolation methods in subsequent subsections using the general linear form (6.4.45). But the reader should remember that all such methods can be applied using local block interpolation (6.4.46) and indeed would need to be for practical use.

#### 6.4.3.2 The "block uniform resampling" (BURS) method

In the context of MRI, Rosenfeld [123] proposed a block interpolation method that he called the "block uniform resampling" (BURS) method.

For simplicity, consider the 1D case with support  $\mathcal{D} = [-FOV/2, FOV/2]$ . It then follows from the sampling theorem that

$$F(\nu) = \sum_{k=-\infty}^{\infty} \operatorname{sinc}(\nu \text{FOV} - k) F(k \text{FOV}).$$

Since f is space limited, its spectrum  $F(\nu)$  must have infinite support, so the above summation must be infinite for exact equality. Nevertheless, if f is approximately band limited, then one could form an approximation using a finite sum in the above expression. This leads to the following relationship:

$$\boldsymbol{v} = \boldsymbol{S} \boldsymbol{u},$$

where we are given  $v_i = F(\nu_i)$ ,  $i = 1, ..., n_d$ , and  $S_{ij} = \operatorname{sinc}(\nu_i \text{FOV} - j)$ , and we want to find the following uniformly spaced samples  $\{u_j = F(\kappa_j)\}$ , where  $\kappa_j = (j - (n_p + 1)/2)/\text{FOV}$ ,  $j = 1, ..., n_p$ . Rosenfeld [123] proposed using the SVD to compute the pseudo inverses of blocks of the rows of S, *i.e.*,

$$\hat{u}_j = \boldsymbol{e}_j' \boldsymbol{S}_j^{\dagger} \boldsymbol{v}_j,$$

where  $v_j$  denotes a subset of v corresponding to nonuniform frequency samples near the *j*th uniform sample location, and  $S_j$  denotes a submatrix of S with the corresponding rows. (And possibly a subset of the columns too.) This is called the **block uniform resampling (BURS)** method. (Similar methods had been proposed previously for image restoration problems, *e.g.*, [158].) Applying rows of the blocks to the data y yields estimates of  $F(\vec{v})$  on a Cartesian grid, and then an inverse FFT is performed to estimate f. This approach was found to be sensitive to noise, so he also proposed regularization methods [159]. Only a limited family of regularization methods is applicable to BURS, whereas iterative methods as described in §6.2.9 can accommodate the entire spectrum of regularization methods and have other advantages as well.

An alternative to the SVD approach is to use the CG algorithm to solve the system of equations  $S'_j S_j u_j = S'_j v_j$ [160]. Another related method is [161].

Another method is [162, 163]. This method can be written as

$$\hat{f} = ZPy$$

where here P is a  $n_d \times n_d$  sparse matrix instead of a diagonal matrix as used in the conjugate phase reconstruction approach (6.4.3). To contain computation, the matrix P is allowed to have a finite number (say  $L \ll n_d$ ) of non-zero elements per row so that the matrix vector multiplication Py is  $O(n_d L)$ . The nonzero elements are selected via an orthogonal matching pursuit (OMP) approach [163]. This method is most useful when the same non-Cartesian sampling pattern will be used for numerous reconstructed images, such as in fMRI or dynamic studies.

#### 6.4.3.3 Min-max spectral interpolation (s,four,freq,minmax)

In the context of band-limited signal interpolation, Yen showed in 1956 that the estimate  $\hat{F}(\vec{\nu})$  that has *minimum* energy  $||F(\cdot)||$  and corresponds to a signal with support  $\mathcal{D}$  has the form (6.4.45) with

$$R_{i}(\vec{\nu}) = \sum_{l=1}^{n_{\rm d}} S(\vec{\nu} - \vec{\nu}_{l}) [\mathbf{K}^{-1}]_{il}, \qquad (6.4.47)^{\text{e, four,}}$$

so that

s,four,freq,minmax

$$\hat{F}(\vec{\nu}) = \sum_{i=1}^{n_{\rm d}} S(\vec{\nu} - \vec{\nu}_i) \left[ \mathbf{K}^{-1} \mathbf{y} \right]_i, \qquad (6.4.48)$$

where  $S(\vec{\nu})$  was defined in (6.1.7), and K is the Fourier crosstalk matrix in (6.1.16) for  $\mathcal{W} = \mathcal{I}$ . This solution is impractical for large  $n_d$ , both due to the large support of  $S(\vec{\nu})$  and due to the large size of K (and K is often ill-conditioned), so block methods are essential for its use in practice. For a non-block method, applying an inverse Fourier transform to (6.4.48) yields the **MNLS** estimator (6.3.4).

This result has been derived from various perspectives [164–167]. Generalizations include consideration of bandpass support [168], multiple dimensions [169], using blocks of neighboring samples [126, 157], multiple dimensions [144, 167], stochastic formulations [152], and using weighting functions other than  $\mathcal{W} = \mathcal{I}$  [126, 170]. Wingham [149] proposed an SVD approach for regularizing the inverse in (6.4.48). Miller regularization [171], where  $\varepsilon I$  is added to the diagonal of K, has also been proposed [149, 170].

It was also found that (6.4.47) is optimal from the following min-max perspective [172]. Defining  $\mathcal{F}_{\mathcal{D}}$  as in (6.4.17), one can show [124] that (6.4.47) is the frequency-domain interpolator that minimizes the worst-case error over  $\mathcal{F}_{\mathcal{D}} \cap \mathcal{N}_{\mathcal{A}}^{\perp}$ . This result is invariant to the norm used [148]. Choi and Munson [124] proposed to replace  $\mathbf{K}^{-1}$  in (6.4.47) with a diagonal matrix  $\mathbf{D}(\mathbf{w})$  where  $\mathbf{w}$  minimizes

Choi and Munson [124] proposed to replace  $K^{-1}$  in (6.4.47) with a diagonal matrix D(w) where w minimizes  $||| K D(w) - I|||_{\text{Frob}}$ . For such a diagonal matrix, the resulting estimator is simply a form of the conjugate phase method, and this criterion for choosing w is a special case of (6.4.41). An interesting open problem would be to find the min-max optimal choice of w with  $S(\cdot)$  in (6.4.47) replaced by a finite-support convolution kernel. Judging from the form of the interpolation error in [124], finding the min-max choice may be challenging.

#### four, freq, iter 6.4.3.4 Iterative spectrum estimation (s, four, freq, iter)

Particularly in the context of band-limited signal interpolation, there have been several proposals for iterative methods for estimating  $F(\vec{\nu})$  from y. Since our ultimate goal is to recover f, probably it is more logical to apply iterative methods to estimate f directly, rather than to first iteratively estimate  $F(\vec{\nu})$  and then compute  $\hat{f}$  by an inverse Fourier transform. That is, unless iterative estimation of  $F(\vec{\nu})$  is less expensive than estimating f. In the interest of completeness we review some of the proposed iterative methods next.

Several authors [173–175] have discussed an iteration for band-limited signal interpolation that is equivalent to the following algorithm for estimating  $F(\vec{\nu})$ :

$$F^{(n+1)}(\vec{\nu}) = F^{(n)}(\vec{\nu}) + \alpha \sum_{i=1}^{n_{\rm d}} S(\vec{\nu} - \vec{\nu}_i) \left[ y_i - F^{(n)}(\vec{\nu}_i) \right], \tag{6.4.49}$$

where  $S(\vec{\nu})$  was defined in (6.1.7). This iteration, known as the **frame algorithm**, is a contraction for certain values of  $\alpha$ , under strong assumptions about the sample locations  $\{\vec{\nu}_i\}$  [173–175]. Taking the inverse Fourier transform and simplifying yields the following equivalent iteration

$$f^{(n+1)} = f^{(n)} + \alpha \mathcal{A}^* (\boldsymbol{y} - \mathcal{A} f^{(n)}). \tag{6.4.50}$$

In this form, it is clear that the iteration is trying to minimize the least-squares cost function ||y - Af||. However, when  $n_d$  is finite, there are a multitude of minimizers of that cost function, so this iteration is inappropriate in that case. For steepest descent and conjugate gradient improvements, see [41, 176].

Other **POCS**-like iterations have also been proposed, *e.g.*, [145]. In the context of the problem considered here, these methods alternate between enforcing "data consistency," *i.e.*,  $y_i = F(\vec{\nu}_i)$ , in the frequency domain, and the space-limited assumption  $f(\vec{x}) = s(\vec{x}) f(\vec{x})$  in the space domain. Specifically, the orthogonal projection onto the convex set of data constraints

$$\mathcal{C}_i \triangleq \{F : F(\vec{\nu}_i) = y_i, F = F * S\}$$

is given by

$$\mathcal{P}_{\mathcal{C}_i} F = F * S + S(\cdot - \vec{\nu}_i) \left[ y_i - (F * S)(\vec{\nu}_i) \right].$$

This closed-form expression leads to a **POCS** algorithm as described in §11.18, although convergence is slow [167].

Aldroubi and Feichtinger [151] propose an iteration for a family of problems that includes our space-limited situation as a special case. In our notation, their iteration has the following form:

$$F^{(n+1)}(\vec{\nu}) = F^{(n)}(\vec{\nu}) + \sum_{i=1}^{n_{d}} S_{i}(\vec{\nu}) [y_{i} - F^{(n)}(\vec{\nu}_{i})]$$

$$F^{(0)}(\vec{\nu}) \triangleq \sum_{i=1}^{n_{d}} y_{i} S_{i}(\vec{\nu})$$

$$S_{i}(\vec{\nu}) \triangleq \sum_{k} \int_{\mathcal{V}_{i}} S(\vec{\nu}' - k/\Delta) \, \mathrm{d}\vec{\nu}' \, S(\vec{\nu} - k/\Delta)$$

$$\mathcal{V}_{i} \triangleq \{\vec{\nu} : \|\vec{\nu} - \vec{\nu}_{i}\| < \|\vec{\nu} - \vec{\nu}_{j}\|, \quad \forall j \neq i\}$$

where  $\Delta$  is the "radius" of  $\mathcal{D}$ , *i.e.*,  $\Delta \triangleq 2 \max_{\vec{x} \in \mathcal{D}} \|\vec{x}\|_{\infty}$ . The set  $\mathcal{V}_i$  is the Voronoi cell associated with the *i*th sample  $\vec{\nu}_i$ . This iteration works by first forming a simple interpolant, and then projecting that interpolant onto the space of spaced-limited functions, and then iterating on the residual error. It is described as a "fast reconstruction procedure" in [151].

Another approach is **deconvolution-interpolation gridding** (**DING**) in which one first iteratively estimates the (smoothed) spectrum of an apodized version of the object, then inverse Fourier transforms and finally de-apodizes [177]. See also [178].

#### s, four, grid 6.4.3.5 Gridding methods (s, four, grid)

Examining the conjugate-phase estimator (6.4.2) in the frequency domain suggests the following frequency-domain estimator:

$$\hat{F}(\vec{\nu}) = \sum_{i=1}^{n_{\rm d}} w_i y_i \, S(\vec{\nu} - \vec{\nu}_i) \,. \tag{6.4.51}$$

This estimator satisfies the convolution property (6.1.6), because S \* S = S since  $s(\vec{x})$  is an indicator function.

The frequency domain estimate (6.4.51) is the foundation of the **gridding** method for reconstruction from Fourier samples. However, since  $S(\vec{\nu})$  is a sinc or sinc-like function that decays slowly, a direct implementation of (6.4.51) would be computationally expensive. In practice, one replaces  $S(\vec{\nu})$  with a "convolution kernel"  $C(\vec{\nu})$  having finite support, suggesting the following estimator initially:

$$\hat{F}_C(\vec{\nu}) = \sum_{i=1}^{n_d} w_i y_i C(\vec{\nu} - \vec{\nu}_i) \,. \tag{6.4.52}$$

Taking the inverse FT, the corresponding estimate in the space domain would be

$$\hat{f}_C(\vec{\mathbf{x}}) = \sum_{i=1}^{n_d} w_i y_i \, \mathrm{e}^{i 2\pi \vec{\nu}_i \cdot \vec{\mathbf{x}}} \, c(\vec{\mathbf{x}}), \quad c(\vec{\mathbf{x}}) \stackrel{\mathrm{FT}}{\longleftrightarrow} C(\vec{\nu}) \,. \tag{6.4.53}$$

Because  $C(\vec{\nu})$  has finite support, its inverse Fourier transform  $c(\vec{x})$  cannot be flat (like  $s(\vec{x})$ ), so one should apply a post-correction for the shape of  $c(\vec{x})$  as follows:

$$\hat{f}(\vec{\mathbf{x}}) = \frac{s(\vec{\mathbf{x}})}{c(\vec{\mathbf{x}})} \hat{f}_C(\vec{\mathbf{x}}) = \frac{s(\vec{\mathbf{x}})}{c(\vec{\mathbf{x}})} \mathcal{F}^{-1} \left\{ \sum_{i=1}^{n_d} w_i y_i \, C(\vec{\nu} - \vec{\nu}_i) \right\} (\vec{\mathbf{x}}), \tag{6.4.54}$$

where  $\mathcal{F}^{-1}$  denotes an inverse  $\bar{d}$ -dimensional Fourier transform. Using the shift-property of the Fourier transform, one can show that this idealized expression simplifies back to the conjugate phase method (6.4.2). However, in practice, rather than using (6.4.54) literally, one implements the transform by sampling  $\hat{F}_C(\vec{\nu})$  on a Cartesian grid and applying an inverse fast Fourier transform (FFT). One can save computation in the sampling step by using precomputed look-up tables [96, 179]. Specifically, letting  $\{\tilde{\nu}_k\}$  denote the Cartesian frequency sample locations, then one evaluates

$$F_k = \hat{F}_C(\tilde{\nu}_k) = \sum_{i=1}^{n_d} w_i y_i \, C(\tilde{\nu}_k - \vec{\nu}_i)$$
(6.4.55)

for k = 0, ..., K - 1. Then one computes the image estimate at spatial sample locations  $\{\vec{x}_n\}$  using an inverse DFT followed by post-correction:

$$\hat{f}(\vec{\mathbf{x}}_n) = \frac{s(\vec{\mathbf{x}}_n)}{c(\vec{\mathbf{x}}_n)} \operatorname{prod}(\Delta_{\vec{\nu}}) \sum_k F_k \, \mathrm{e}^{i2\pi \vec{\mathbf{x}}_n \cdot \tilde{\boldsymbol{\nu}}_k} \,. \tag{6.4.56}$$

This inverse DFT is amenable to an inverse FFT provided one chooses the sample locations such that  $\vec{x}_n \cdot \tilde{\nu}_k = \vec{n} \cdot \vec{k} / K$ .

For analytical insight, consider the continuous-space estimate corresponding to (6.4.56), in the 1D case and for an infinite number of Cartesian frequency samples:

$$\hat{f}(x) = \frac{s(x)}{c(x)} \bigtriangleup_{\nu} \sum_{k=-\infty}^{\infty} F_k e^{i2\pi x k \bigtriangleup_{\nu}} = \frac{s(x)}{c(x)} \sum_{i=1}^{n_d} w_i y_i e^{i2\pi\nu_i x} \bigtriangleup_{\nu} \sum_k C(k \bigtriangleup_{\nu} - \nu_i) e^{i2\pi x (k\bigtriangleup_{\nu} - \nu_i)}$$
$$= \frac{s(x)}{c(x)} \sum_{i=1}^{n_d} w_i y_i e^{i2\pi\nu_i x} \sum_{l=-\infty}^{\infty} e^{-i2\pi\nu_i l/\bigtriangleup_{\nu}} c(x - l/\bigtriangleup_{\nu}).$$
(6.4.57)

In words: sampling in the Fourier domain causes replication of c(x) in the spatial domain.

One would like to choose a kernel  $C(\vec{v})$  that closely approximates  $S(\vec{v})$  yet has finite support, and for which  $c(\vec{x})$  closely approximates  $s(\vec{x})$ , *i.e.*, is nearly flat over  $\mathcal{D}$  and otherwise zero. These are conflicting goals. In practice, Kaiser-Bessel functions have been found to be a useful compromise [16]. The importance of the choice of this kernel depends on the Cartesian sampling used. If one uses substantial over sampling, then one may be able to use simple nearest-neighbor interpolation (*i.e.*, a rectangular kernel) at the price of more expensive FFT computations [180, 181]. This may be a reasonable trade-off in some cases since optimized hardware may be more readily available for computing 2D FFTs than for the type of convolution required in (6.4.52). Conversely, if minimal over sampling is used, then more care is needed in choosing the interpolator. In particular, Beatty et al. [179] recommend using a Kaiser-Bessel interpolator, *cf.* (6.6.18), with order m = 0 and shape parameter

$$\alpha = \pi \sqrt{\frac{J^2}{(K/N)^2} \left(\frac{K}{N} - \frac{1}{2}\right)^2 - 0.8},$$
(6.4.58)

where  $K/N \in [1.25, 2]$  denotes the FFT over-sampling factor and J denotes the width of the support of  $C(\nu)$  relative to  $\Delta_{\nu}$ . This choice (approximately) minimizes the **aliasing amplitude** [179, eqn. (5)].

#### four, freq, impl 6.4.3.6 Practical gridding implementation (s, four, freq, impl)

In practice, one implements the frequency-domain gridding formula (6.4.52) using a finite set of Cartesian frequency sample locations as described in (6.4.55). There are severals ways one can implement this type of operation, illustrated in Fig. 6.4.2. To describe these methods mathematically, let  $C \subset \mathbb{R}^{n_d}$  denote the support of  $C(\vec{\nu})$ .

#### • Pulling from a neighborhood (aka "output driven")

Because  $C(\vec{v})$  has finite support C, the summation in (6.4.52) only involves a few nonzero terms. In particular, defining

$$\mathcal{I}_k \triangleq \{i : \tilde{\nu}_k - \vec{\nu}_i \in \mathcal{C}\}$$

one can implement gridding by looping over each desired Cartesian location  $\tilde{\nu}_k$  and "pulling" values from the neighbors in  $\mathcal{I}_k$  as follows:

for k = 1, ..., K:

$$F_k = \sum_{i \in \mathcal{I}_k} w_i y_i \, C(\tilde{\nu}_k - \vec{\nu}_i) \, .$$

A disadvantage of this approach is that it requires determination of the set  $\mathcal{I}_k$  which is inconvenient when the  $\vec{\nu}_i$  values are arbitrary. See [182] for a GPU version and [140] for a helpful compartmentalization scheme; however, a **k-d tree** is likely to be more efficient than the partition in [140].

#### • Pushing to a neighborhood (aka "input driven")

An alternative approach is to move *i* to the outer loop and *k* to the inner loop. For each frequency sample location  $\vec{\nu}_i$ , we "push" the (weighted) value  $y_i$  to the appropriate Cartesian neighbors, defined by the set

$$\mathcal{K}_i \triangleq \{k : \tilde{\nu}_k - \vec{\nu}_i \in \mathcal{C}\}$$

One can implement this approach as follows.

Initialize 
$$F_k = 0, \ k = 1, \dots, K$$
  
for  $i = 1, \dots, n_d$ :  
for  $k \in \mathcal{K}_i$ :  
 $F_k += w_i y_i C(\tilde{\nu}_k - \vec{\nu}_i)$ . (6.4.59)

Mathematically, both pulling from and pushing to a neighborhood yield the same result, but pushing is easier to implement because the set *K<sub>i</sub>* is simple to form due to the regular spacing of the Cartesian sample locations {*ν<sub>k</sub>*}.
Pulling from *L* nearest neighbors

Yet another option is to pull from the L nearest neighbors of each Cartesian point, defined by

 $\mathcal{N}_k \triangleq \{i_1, \ldots, i_L : \vec{\nu}_{i_l} \text{ are closer to } \tilde{\nu}_k \text{ than the others} \}.$ 

The implementation would be

$$F_k = \frac{\sum_{i \in \mathcal{N}_k} w_i y_i C(\tilde{\nu}_k - \vec{\nu}_i)}{\sum_{i \in \mathcal{N}_k} w_i C(\tilde{\nu}_k - \vec{\nu}_i)}$$

In this case, the interpolation operation is *not* equivalent to a convolution, so the post-correction in (6.4.56) is inapplicable, so one should normalize the interpolator in the frequency domain as shown above.

One can also devise a pushing implementation of this L nearest neighbor approach, which requires two arrays, one for accumulating the numerators and one for the denominators above.

Assessing the relative merits of these approaches is apparently an *open problem*. Historically the pushing approach (6.4.59) seemed the most convenient for practical implementation [111], but parallel computing (*e.g.*, GPU) can favor pulling [182].

## 6.5 Summary (s,four,summ)

This *chapter* has surveyed a variety of methods for image reconstruction from Fourier-domain samples. The gridding method (§6.4.3.5) is among the fastest methods, and is thus used frequently for well-sampled data. However, gridding requires **density compensation factors**, and there are many proposed approaches for choosing those factors. Iterative methods do not require density compensation factors. Of such methods, the preconditioned conjugate gradient (PCG) iteration (§6.2.9) using FFTs to perform Toeplitz matrix multiplication (§6.2.10) is particularly efficient for quadratic regularization.



Figure 6.4.2: Practical implementation of methods for gridding from non-Cartesian frequency samples (blue circles)<sub>g\_four\_interp2\_grid</sub> to a Cartesian grid (red squares).

Although resolution noise trade-offs have been analyzed for density-weighted conjugate-phase reconstructions in MR [183], such analyses are an *open problem* for more general reconstruction approaches such as the regularized methods described here.

Many of the methods discussed here for finding density compensation factors appear impractical as proposed initially, but some may become feasible if combined with the "block" approaches of §6.4.3.1. See [124] as an example.

Although the methods discussed herein are applicable to arbitrary sampling patterns, the quality of the results will of course depend greatly on the sampling pattern. See [184] for discussion of **stability** of recovery in terms of the sampling pattern.

## 6.6 Appendix: NUFFT calculations (s,four,nufft)

For equally spaced basis functions, the matrix-vector multiplication Ax needed in (6.2.16) is

$$[\mathbf{A}\mathbf{x}]_i = \sum_{j=1}^{n_{\rm p}} a_{ij} x_j = B(\vec{\nu}_i) \sum_{j=1}^{n_{\rm p}} x_j \, \mathrm{e}^{-i2\pi\vec{\nu}_i \cdot \vec{\mathbf{x}}_j} \,,$$

using (6.2.12). This operation involves a type of **non-uniform FFT** (**NUFFT**) calculation [50]. This section reviews the 1D NUFFT; the extension to higher dimensions is fairly easy.

#### MIRT NUFFT software is available.

Consider the problem of computing the discrete-time Fourier transform (DTFT) of a discrete-time signal x[n], n = 0, ..., N - 1, at (unequally spaced) frequency samples  $\omega_m, m = 1, ..., M$ , as follows:

$$X(\omega_m) = \sum_{n=0}^{N-1} x[n] e^{-i\omega_m n} .$$
 (6.6.1)

Directly evaluating (6.6.1) would require MN operations, or  $O(N^2)$  in the usual case where  $M \approx N$ . NUFFT methods require roughly  $O(N \log N)$  operations and work as follows.

- Multiply x[n] by some nonzero scaling factors s[n].
- Compute an over-sampled DFT of the scaled signal.
- Interpolate the DFT values onto the desired frequency locations.

Mathematically, the over-sampled (K-point) DFT is given by

$$Y[k] \triangleq \sum_{n=0}^{N-1} s[n] x[n] e^{-i\gamma kn}, \qquad k = 0, \dots, K-1,$$
(6.6.2)

where  $K \ge N$  and  $\gamma \triangleq 2\pi/K$ . Typically one uses  $N < K \le 2N$  [179]. Now the goal is to compute  $X(\omega_m)$  (usually approximately) from the DFT values  $\{Y[k]\}$  via interpolation.

#### 6.6.1 Basic ideal interpolator

It is instructive to first consider a simple ideal interpolator for calculating  $X(\omega_m)$  from  $\{Y[k]\}$ . Inverting (6.6.2) and substituting into (6.6.1) and simplifying yields

$$X(\omega_m) = \sum_{n=0}^{N-1} \left( \frac{1}{s[n]} \frac{1}{K} \sum_{k=0}^{K-1} Y[k] e^{i\gamma kn} \right) e^{-i\omega_m n} = \sum_{k=0}^{K-1} Y[k] \tilde{I}_c \left( \frac{\omega_m}{\gamma} - k \right), \tag{6.6.3}$$

where an ideal interpolator (for given scaling factors s[n]) is the following K-periodic complex-valued function:

$$\tilde{I}_{c}(\kappa) \triangleq \frac{1}{K} \sum_{n=0}^{N-1} \frac{1}{s[n]} e^{-i\gamma\kappa n}, \quad \kappa \in \mathbb{R}.$$
(6.6.4)

Usually one chooses the scaling factors s[n] to be symmetric in the following sense: s[N-1-n] = s[n], n = 0, ..., N-1, in which case the ideal interpolator simplifies (see Problem 6.26) to the product of a linear phase term and a real, symmetric kernel:

$$\tilde{I}_{c}(\kappa) = \frac{N}{K} e^{-i\gamma\kappa\eta_{0}} I_{r}(\kappa), \qquad I_{r}(\kappa) \triangleq \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{s[n]} e^{-i\gamma\kappa(n-\eta_{0})} = \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{s[n]} \cos(\gamma\kappa(n-\eta_{0})), \qquad (6.6.5)$$

where  $\eta_0 \triangleq (N-1)/2$  denotes the "midpoint" of the signal sample locations.

In particular, if s[n] = 1, then the interpolator is  $I_r(\kappa) = \delta_{N,K}(\kappa)$ , where  $\delta_{N,K}(\cdot)$  denotes the **Dirichlet kernel**:

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{-i\gamma\kappa(n-\eta_0)} = \frac{\sin(\pi\kappa N/K)}{N\sin(\pi\kappa/K)} \triangleq \delta_{N,K}(\kappa).$$
(6.6.6)

This function is 2K-periodic when N is even and is K-periodic when N is odd. See Fig. 6.6.1.

The function  $I_c$  is K-periodic, whereas the functions  $e^{-i\gamma\kappa\eta_0}$  and  $I_r(\kappa)$  are 2K-periodic in the usual case where N is even. Because Y[k] is K-periodic, it seems more natural to work with K-periodic interpolators. The next section addresses this issue.

#### 6.6.2 Generalized ideal interpolator

In the usual case where K > N, the fact that the sum in (6.6.2) involves only n = 0, ..., N - 1 means that there are many ideal interpolators! We generalize (6.6.3) as follows:

$$X(\omega_m) = \sum_{n=0}^{K-1} \left( q[n] \frac{1}{K} \sum_{k=0}^{K-1} Y[k] e^{i\gamma kn} \right) e^{-i\omega_m n} = \sum_{k=0}^{K-1} Y[k] \tilde{I}_c \left( \frac{\omega_m}{\gamma} - k \right), \tag{6.6.7}$$

where  $q[n] = \begin{cases} \frac{1}{s[n]}, & n = 0, \dots, N-1 \\ arbitrary in \mathbb{R}, & n = N, \dots, K-1. \end{cases}$  For any given choice of q[n] factors there is a corresponding *K*-periodic ideal interpolation kernel defined similarly to (6.6.4) as follows:

$$\tilde{I}_{\rm c}(\kappa) \triangleq \frac{1}{K} \sum_{n=0}^{K-1} q[n] \,\mathrm{e}^{-\imath \gamma \kappa n} \,. \tag{6.6.8}$$

We have considerable freedom in designing the q[n] factors, so it seems natural to make a choice that avoids the issues with 2K-periodicity for even N that arose above.

A natural approach is to choose some M satisfying  $N \le M \le K - 1$  and select q[n] factors having the following symmetry property:  $q[n] = \begin{cases} q[M-1-n], & n = 0, \dots, M-1 \\ 0, & \text{otherwise.} \end{cases}$  In such cases, the generalized ideal interpolator (6.6.8) simplifies to the product of a linear phase term and a real, symmetric kernel as follows (see Problem 6.26):

$$\tilde{I}_{c}(\kappa) = \frac{M}{K} e^{-i\gamma\kappa\eta_{M}} \tilde{I}_{r}(\kappa)$$
(6.6.9)

$$\tilde{I}_{\rm r}(\kappa) \triangleq \frac{1}{M} \sum_{n=0}^{M-1} q[n] \,{\rm e}^{-\imath \gamma \kappa (n-\eta_M)} = \frac{1}{M} \sum_{n=0}^{M-1} q[n] \cos(\gamma \kappa (n-\eta_M)), \tag{6.6.10}$$

where  $\eta_M \triangleq (M-1)/2$ . If we choose  $q[n] = 1, n = 0, \dots, M-1$ , then

$$\tilde{I}_{\rm r}(\kappa) = \frac{\sin(\pi\kappa M/K)}{M\sin(\pi\kappa/K)} = \delta_{M,K}(\kappa) \,.$$

In particular, as long as we always choose M to be odd, which is always possible when N < K, then both the phase factor  $e^{-i\gamma\kappa\eta_M}$  and the interpolator kernel  $\tilde{I}_r(\kappa)$  will be K-periodic. See Fig. 6.6.1.

Using the factorization in (6.6.10), we rewrite the interpolation expression (6.6.7) as

$$X(\omega_m) = e^{-\iota\omega_m \eta_M} \bar{X}(\omega_m) \tag{6.6.11}$$

$$\bar{X}(\omega_m) \triangleq \sum_{k=0}^{K-1} \bar{Y}[k] \frac{M}{K} \tilde{I}_r \left(\frac{\omega_m}{\gamma} - k\right)$$
(6.6.12)

$$\bar{Y}[k] \triangleq e^{i\gamma k\eta_M} Y[k] \,. \tag{6.6.13}$$

In this form we see that the ideal interpolator is essentially real; pre-multiplying Y[k] by  $e^{i\gamma k\eta_M}$  and post-multiplying by  $e^{-i\omega_m\eta_M}$  simply account for the limits of the sums in (6.6.1) and (6.6.2); if both of those sums were shifted by  $\eta_M$  then the phase factors would disappear entirely. (This simplification was overlooked in [50].)

How might we use effectively the freedom in choosing the q[n] factors? Anticipating the practical interpolators in the next section, one option is to try to minimize the energy in the kernel outside some interval:  $\int_{J/2}^{K/2} \tilde{I}_{r}^{2}(\kappa) d\kappa$ . For example, we could maximize  $\int_{0}^{J/2} \tilde{I}_{r}^{2}(\kappa) d\kappa / \int_{0}^{K/2} \tilde{I}_{r}^{2}(\kappa) d\kappa$ . This problem is related to **prolate spheroidal wave functions** [185, 186] that seem inconvenient for practical use.

#### 6.6.3 Practical interpolators

An ideal interpolator (6.6.13) or (6.6.7) would require KM operations, which is impractical. In practice, we replace the interpolator in the interpolation step (6.6.13) by an approximation:

$$\bar{X}(\omega_m) \approx \sum_{k=0}^{K-1} \bar{Y}[k] \,\tilde{\psi}\left(\frac{\omega_m}{\gamma} - k\right),\tag{6.6.14}$$



Figure 6.6.1: Comparison of ideal real interpolation kernels  $\tilde{I}_r$  for N = 8, K = 16 and  $M \in \{8, 13\}$ .

where  $\tilde{\psi}(\cdot)$  is the *K*-periodic extension of a *finite-support* interpolation kernel  $\psi(\cdot)$ , *i.e.*,

$$\tilde{\psi}(\kappa) \triangleq \sum_{l=-\infty}^{\infty} \psi(\kappa - lK).$$

If the kernel  $\psi(\cdot)$  has support [-J/2, J/2), then the summation (6.6.14) involves only J terms, *i.e.*,

$$\sum_{k=0}^{K-1} \bar{Y}[k] \,\tilde{\psi}\left(\frac{\omega_m}{\gamma} - k\right) = \sum_{j=1}^{J} \bar{Y}[(k_0(\omega_m) + j) \operatorname{mod} K] \,\psi\left(\frac{\omega_m}{\gamma} - [k_0(\omega_m) + j]\right), \tag{6.6.15}$$

where<sup>7</sup>

$$k_0(\omega) \triangleq \left\lfloor \frac{\omega}{\gamma} - \frac{J}{2} \right\rfloor, \tag{6.6.16}$$

where  $|\cdot|$  denotes the integer floor function. This offset satisfies the following (integer) shift invariance property:

$$k_0(\omega + l\gamma) = l + k_0(\omega), \qquad \forall l \in \mathbb{Z}.$$
(6.6.17)

The interpolation (6.6.15) uses only the J nearest DFT neighbors, reducing the computation to JM operations, where  $J \ll K$ .

A good choice for the kernel is the generalized Kaiser-Bessel function [188, 189]:

$$\psi(\kappa) = f_J^m(\kappa) \frac{I_m(\alpha f_J(\kappa))}{I_m(\alpha)}, \tag{6.6.18}$$

where  $I_m$  denotes the modified Bessel function of order m, and

$$f_J(\kappa) \triangleq \begin{cases} \sqrt{1 - \left(\frac{\kappa}{J/2}\right)^2}, & |\kappa| < J/2 \\ 0, & \text{otherwise.} \end{cases}$$

The shape of this function is related to the "shape parameter"  $\alpha$ . Good choices are m = 0 and  $\alpha = 2.34J$  for K/N = 2 [50]. See also (6.4.58) from [179, eqn. (5)], and the formula recommended by Wajer et al. [190]. For the Kaiser-Bessel interpolator, the corresponding scaling factors are

$$s[n] = 1/\Psi((n - \eta_0)/K), \tag{6.6.19}$$

where [188]:

$$\Psi(x) = (1/2)^m \pi^{\bar{d}/2} (J/2)^{\bar{d}} \alpha^m \Lambda(z(x)) / I_m(\alpha) \stackrel{\text{FT}}{\longleftrightarrow} \psi(u), \qquad (6.6.20)$$

where  $\bar{d} = 1$  (for 1D case),  $\nu = \bar{d}/2 + m$ ,  $z(x) = \sqrt{(\pi J x)^2 - \alpha^2}$ , and  $\Lambda(z) = (z/2)^{-\nu} J_{\nu}(z)$ , where  $J_{\nu}$  denotes the Bessel function of the first kind of order  $\nu$ . See [191] for an approach to optimize the scaling factors s[n]. See also [192] for a mean-square optimal NUFFT approach.

<sup>7</sup> The following expressions were given in [50] and [187] respectively:  $k_0(\omega) \triangleq \begin{cases} (\arg\min_{k \in \mathbb{Z}} |\omega/\gamma - k|) - \frac{J+1}{2}, & J \text{ odd} \\ (\max\{k \in \mathbb{Z} : \omega/\gamma \ge k\}) - \frac{J}{2}, & J \text{ even}, \end{cases}$ 

and  $k_0(\omega) \triangleq \begin{cases} [\omega/\gamma] - \frac{J+1}{2}, & J \text{ odd} \\ [\omega/\gamma] - \frac{J}{2}, & J \text{ even,} \end{cases}$  where  $[\cdot]$  denotes rounding to the nearest integer. Both of these simplify to (6.6.16).

Combining the interpolation and the over-sampled DFT, we have

$$X(\omega_m) \approx \sum_{k=0}^{K-1} e^{-i\gamma(\omega_m/\gamma - k)\eta_0} \tilde{\psi}\left(\frac{\omega_m}{\gamma} - k\right) \left[\sum_{n=0}^{N-1} s[n] x[n] e^{-i\gamma kn}\right]$$
$$= \sum_{n=0}^{N-1} x[n] \left[s[n] \sum_{k=0}^{K-1} e^{-i\gamma(\omega_m/\gamma - k)\eta_0} \tilde{\psi}\left(\frac{\omega_m}{\gamma} - k\right) e^{-i\gamma kn}\right]$$

This requires  $O(K \log K)$  operations for the FFT step and JM operations for the interpolation step. Comparing this expression with the original DTFT expression (6.6.1) reveals the following underlying approximation:

$$e^{-\iota\omega(n-\eta_0)} \approx s[n] \sum_{k=0}^{K-1} \tilde{\psi}\left(\frac{\omega}{\gamma} - k\right) e^{-\iota\gamma k(n-\eta_0)}.$$
(6.6.21)

To evaluate the gradient in (6.2.16), we also need the ability to perform the adjoint operation A'y, where

$$[\mathbf{A}'\mathbf{y}]_j = \sum_{i=1}^{n_{\rm d}} a_{ij}^* y_i = \sum_{i=1}^{n_{\rm d}} B(\vec{\nu}_i) e^{i2\pi\vec{\nu}_i \cdot \vec{x}_j} y_i.$$

This is an "adjoint NUFFT" operation. In 1D, the equivalent expression in signal processing notation is

$$\bar{x}[n] = \sum_{m=1}^{M} X_m e^{i\omega_m n}$$

Using the complex conjugate of the approximation (6.6.21), we have the following **NUFFT adjoint** operation:

$$\bar{x}[n] \approx s^*[n] \sum_{k=0}^{K-1} \left[ \sum_{m=1}^M \bar{X}_m \,\tilde{\psi} \left( \frac{\omega_m}{\gamma} - k \right) \right] \mathrm{e}^{-\imath \gamma k \eta_0} \, \mathrm{e}^{\imath \gamma k n} \,, \tag{6.6.22}$$

where  $\bar{X}_m \triangleq e^{i\omega_m \eta_0} X_m$ . The bracketed expression is called **gridding**, and the outer summation is an inverse DFT (to within a scale factor).

MIRT | The approximation (6.6.22) is implemented in nufft\_adj.m and newfft.m.

The extension from 1D to 2D and higher is relatively straightforward [50]. For image reconstruction purposes, each  $\vec{\omega}_m$  is the product of  $2\pi \vec{\nu}_i$  and the pixel size.

MIRT The Gnufft object encapsulates these operations so that one can type A \* x and A' \* y to invoke (6.6.14) and (6.6.22).

## 6.7 Appendix: NUFFT-based gridding (s,four,nufft,grid)

The section describes details about how to implement the conjugate phase method (6.4.2). For simplicity we describe a 1D version of the problem. One would like to evaluate summations of the form<sup>8</sup>

$$f[n] = \sum_{m=1}^{M} F_m e^{i\omega_m n}, \qquad n = 0, \dots, N-1.$$
(6.7.1)

This is the "type 1" NUFFT as defined for example in [54]. One can assume without loss of generality, that  $|\omega_m| \le \pi$ . The central approximation that underlies all gridding methods is the following:

$$e^{i\omega n} \approx e_n(\omega) \triangleq s[n] \sum_{k=0}^{K-1} \tilde{\psi}_k(\omega/\gamma) e^{i\gamma kn}, \qquad (6.7.2)$$

for some  $K \ge N$ , where  $\gamma \triangleq 2\pi/K$  and  $\left\{\tilde{\psi}_k(\omega/\gamma)\right\}_{k=0}^{K-1}$  denotes interpolation coefficients associated with frequency  $\omega$ . The (positive) s[n] values are called scaling factors [105]. The design problem is to choose the scaling factors  $s = (s[0], \ldots, s[N-1])$  and the interpolator  $\tilde{\psi}_k$  to minimize the approximation error in (6.7.2).

<sup>&</sup>lt;sup>8</sup> In MRI, a summation of the form  $\sum_{m=1}^{M} F_m e^{i\omega_m (n-N/2)}$  may be needed, but this is equivalent to (6.7.1) by modifying  $F_m$ .

If one imposed no constraints on  $\tilde{\psi}_k$ , then (6.7.2) could be made to be exact by using the following ideal interpolator:

$$\tilde{\psi}_{k}(\kappa) = \frac{1}{K} \sum_{n=0}^{N-1} \frac{1}{s[n]} e^{i\gamma\kappa n} e^{-i\gamma kn} = \frac{N}{K} \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{s[n]} e^{i\gamma(\kappa-k)n} = \tilde{\psi}_{0}(\kappa-k), \qquad (6.7.3)$$

for k = 0, ..., K - 1, where

$$\tilde{\psi}_0(\kappa) \triangleq \frac{N}{K} e^{-i\gamma\kappa\eta_0} \,\delta_{N,K}(\kappa), \tag{6.7.4}$$

where  $\eta_0 \triangleq (N-1)/2$  and  $\delta_{N,K}$  denotes the following K-periodic Dirichlet kernel:

$$\delta_{N,K}(\kappa) \triangleq \frac{1}{N} \sum_{n=0}^{N-1} e^{i\gamma\kappa(n-\eta_0)} = \begin{cases} \frac{\sin(\pi\kappa N/K)}{N\sin(\pi\kappa/K)}, & \kappa/K \notin \mathbb{Z} \\ 1, & \kappa/K \in \mathbb{Z}. \end{cases}$$
(6.7.5)

Note that  $\delta_{N,K}(\kappa)$ , which is the "core" of the interpolator is real; the only complex aspect of (6.7.4) is the linear phase term. Furthermore, the  $\kappa - k$  argument in (6.7.3) indicates an (integer) shift invariance. It is logical therefore to expect that other useful interpolators will have such integer shift invariance and will have the same form as (6.7.4), namely, the product of a linear phase term with a real interpolation kernel.

The ideal interpolator (6.7.4) would be computationally impractical, so in practice one allows  $\tilde{\psi}_k(\kappa)$  to have at most  $J \ll K$  nonzero values for each  $\omega$ . In particular, for simplicity one uses the J neighbors that are nearest to  $\omega$  (in a modulo- $2\pi$  sense). Using the integer offset defined in , we replace (6.7.2) by the equivalent expression

$$e^{i\omega n} \approx \hat{e}_n(\omega; \boldsymbol{u}) = s[n] \sum_{j=1}^J u_j(\omega) e^{i\gamma(k_0(\omega) + j)n}, \qquad (6.7.6)$$

where  $u(\omega) = (u_1(\omega), \ldots, u_J(\omega))$  denotes the vector of length J of interpolation coefficients associated with frequency  $\omega$ .

In principle the interpolation coefficients could be shift variant, but we will show shortly that u satisfies a shiftinvariance property equivalent to that of (6.7.3), contrary to the implication of [193, Fig. 1]. Having made the approximation (6.7.6) we can evaluate the NUFFT (approximately) as follows:

$$f[n] \approx \hat{f}[n] \triangleq \sum_{m=1}^{M} F_m \left[ s[n] \sum_{j=1}^{J} u_j(\omega_m) e^{i\gamma(k_0(\omega_m)+j)n} \right]$$
$$= s[n] \sum_{j=1}^{J} \left[ \sum_{m=1}^{M} F_m u_j(\omega_m) e^{i\gamma(k_0(\omega_m)+j)n} \right]$$
$$= s[n] \sum_{k=0}^{K-1} \left[ \sum_{m=1}^{M} F_m u_{k-k_0(\omega_m)}(\omega_m) \mathbb{I}_{\{1 \le k-k_0(\omega_m) \le J\}} \right] e^{i\gamma kn}$$

The inner summation is **gridding** from the nonuniform frequency space sample locations  $\{\omega_m\}$  onto the nearby Cartesian samples  $\{\gamma k\}$ , and the outer summation is an inverse FFT.

The approximation (6.7.2) leads to the following bounds on the average weighted squared error:

$$\frac{1}{N} \sum_{n=0}^{N-1} w[n] \left| f[n] - \hat{f}[n] \right|^2 = \frac{1}{N} \sum_{n=0}^{N-1} w[n] \left| \sum_{m=1}^M F_m \left( e^{i\omega_m n} - \hat{e}_n(\omega_m; \boldsymbol{u}) \right) \right|^2$$
$$\leq \frac{1}{N} \sum_{n=0}^{N-1} w[n] \sum_{m=1}^M |F_m|^2 \left| e^{i\omega_m n} - \hat{e}_n(\omega_m; \boldsymbol{u}) \right|^2$$
$$\leq \left( \sum_{m=1}^M |F_m|^2 \right) \max_{\omega} E^2(\omega; \boldsymbol{s}; \boldsymbol{u}),$$

where

$$E^{2}(\omega; \boldsymbol{s}; \boldsymbol{u}) \triangleq \frac{1}{N} \sum_{n=0}^{N-1} w[n] \left| e^{i\omega n} - \hat{e}_{n}(\omega; \boldsymbol{u}) \right|^{2}.$$

$$(6.7.7)$$

This upper bound is tight because it would be achieved for a spectrum  $F_m$  that concentrates entirely on the worst-case frequency. In light of this tight bound, it is desirable to design the scaling factors s and the interpolation coefficients  $u(\omega)$  to minimize the worst-case error by the following min-max criterion (cf. [50, eqn. (10)]):

$$\min_{\boldsymbol{s}\in\mathbb{C}^N} \max_{\boldsymbol{\omega}} \min_{\boldsymbol{u}\in\mathbb{C}^J} E(\boldsymbol{\omega};\boldsymbol{s};\boldsymbol{u}).$$
(6.7.8)

In particular, for a given choice of s, we design  $u(\omega)$  for each  $\omega$  as follows

$$\boldsymbol{u}(\omega) \triangleq \mathop{\arg\min}_{\boldsymbol{u} \in \mathbb{C}^J} E(\omega; \boldsymbol{s}; \boldsymbol{u})$$

Having optimized  $u(\omega)$ , the worst-case error is the following:

$$E_{\max}(\boldsymbol{s}) \triangleq \max_{\omega} E(\omega; \boldsymbol{s}; \boldsymbol{u}(\omega)).$$
(6.7.9)

Finding  $u(\omega)$  by minimizing (6.7.7) is simply a weighted least-squares problem that is linear in u. Solving that minimization directly leads to complex expressions, *e.g.*, [193, eqn. (7)] that obscure the nature of the interpolator. Instead we rewrite the error as follows (*cf.* [50, eqn. (14)]:

$$E^{2}(\omega;\boldsymbol{s};\boldsymbol{u}) = \sum_{n=0}^{N-1} w[n] \left| e^{i\omega n} - s[n] \sum_{j=1}^{J} u_{j}(\omega) e^{i\gamma(k_{0}(\omega)+j)n} \right|^{2} = \left\| \boldsymbol{b} - \boldsymbol{SC} \boldsymbol{\Lambda}(\omega) \boldsymbol{u}(\omega) \right\|_{\boldsymbol{W}^{1/2}}^{2}$$

where

$$\begin{split} \boldsymbol{W} &\triangleq \frac{1}{N} \operatorname{diag}\{\boldsymbol{w}[n]\} \\ \boldsymbol{b}_{n}(\omega) &\triangleq \mathrm{e}^{i(\omega-\gamma k_{0}(\omega))(n-\eta_{0})} \\ \boldsymbol{S} &\triangleq \operatorname{diag}\{\boldsymbol{s}[n]\} \\ \boldsymbol{C}_{nj} &\triangleq \mathrm{e}^{i\gamma j(n-\eta_{0})} \\ \boldsymbol{\Lambda}_{ij}(\omega) &\triangleq \mathrm{e}^{-i(\omega-\gamma (k_{0}(\omega)+j))\eta_{0}} \,. \end{split}$$
(6.7.10)

Clearly the (weighted) LS minimizer is

$$\boldsymbol{u}(\omega) = \boldsymbol{\Lambda}^*(\omega) \, \boldsymbol{T}^{-1} \, \boldsymbol{r}(\omega),$$

because  $\Lambda^{-1} = \Lambda^*,$  where

$$oldsymbol{T} riangleq oldsymbol{C}' oldsymbol{S}' WSC$$
  
 $oldsymbol{r}(\omega) riangleq oldsymbol{C}' oldsymbol{S}' Wb(\omega).$ 

The elements of the  $J \times J$  Toeplitz matrix T and the N-vector  $r(\omega)$  are given by

$$T_{lj} = \frac{1}{N} \sum_{n=0}^{N-1} w[n] s^{2}[n] e^{i\gamma(l-j)(n-\eta_{0})}$$
$$r_{j}(\omega) = \frac{1}{N} \sum_{n=0}^{N-1} w[n] s[n] e^{i(\omega-\gamma(k_{0}(\omega)+j))(n-\eta_{0})}$$

The key to this representation is the following fact: in the usual case where w[n] and s[n] are chosen to be symmetric about  $\eta_0$ , then T is a *real* matrix, and  $r(\omega)$  is a *real* vector. In particular, the elements simplify to

$$T_{lj} = \frac{1}{N} \sum_{n=0}^{N-1} w[n] s^{2}[n] \cos(\gamma(l-j)(n-\eta_{0}))$$
$$r_{j}(\omega) = \frac{1}{N} \sum_{n=0}^{N-1} w[n] s[n] \cos((\omega - \gamma (k_{0}(\omega) + j))(n-\eta_{0})).$$

Thus the core of this LS optimal interpolator is real; the only complex aspects are the linear phase factors in (6.7.11), and these correspond directly to the phase of the ideal interpolator in (6.7.4). Furthermore, the frequency  $\omega$  enters

the expressions above only in the form  $\omega - \gamma k_0(\omega)$ , so this LS optimal interpolator also satisfies the integer shift invariance seen in the ideal interpolator (6.7.4).

Having optimized the coefficients  $u(\omega)$ , the next step is to design s to minimize the worst-case error in (6.7.9). Unfortunately there is no apparent analytical optimizer for s. Therefore, one must use numerical methods to optimize s. In the literature, several choices for s have been proposed, all of which have the form (6.6.19) for various choices for the function  $\Psi(t)$ , such as the following:

- uniform factors, *i.e.*,  $\Psi(t) = 1$ ,
- cosine factors [105]:  $\Psi(t) = \cos(\pi t)$ ,
- Gaussian factors [98]:  $\Psi(t) = \sigma \sqrt{2\pi} e^{-\pi (t\sigma \sqrt{2\pi})^2}$ , with  $\sigma$  chosen to minimize  $E_{\text{max}}$ ,
- and Kaiser-Bessel factors (6.6.20), with parameters optimized per (6.4.58) [179, eqn. (5)].

Fig. 6.7.1 shows the error  $E(\omega, s, u(\omega))$  for each of the above choices of the scaling factors. The worst-case error is smallest for the Kaiser-Bessel scaling factors.



fig\_nufft\_emax\_err\_

Figure 6.7.1: Plots of the error  $E(\omega, s, u(\omega))$  vs  $\omega/\gamma$  for various choices of the scaling factors s.

The results in Fig. 6.7.1 were for the unweighted error criterion with w[n] = 1. This criterion seems the most natural because usually we lack prior information that would favor weighting the error in some parts of the field of view more or less than other parts. However, note that if we choose  $w[n] = 1/s^2[n]$ , then the matrix T becomes independent of s, and leads to a simple closed-form solution [50, eqn. (29)]:

$$T_{lj} = \delta_{N,K}(j-l),$$

where  $\delta_{N,K}$  was defined in (6.7.5). Furthermore, if we expand 1/s[n] in terms of a suitable Fourier series, then a closed-form solution for  $r_j(\omega)$  is also available [50, eqn. (30)]. However, because  $u(\omega)$  can be tabulated easily for a finely sampled grid of  $\omega$  values, closed-form expressions for T and  $r(\omega)$  are unessential. The convenience of such expressions is offset by the subjectivity in choosing w[n] values that depend on s.



Figure 6.7.2: Interpolator cores

## 6.8 Appendix: Toeplitz matrix-vector multiplication (s,four,toep)

This section summarizes how to perform matrix-vector multiplication Tx when T is a Toeplitz matrix, using FFT operations, focusing on the 2D case. Computing y = Tx, where T is Toeplitz, is equivalent to the following operation in signal processing notation:

$$y[m,n] = \sum_{m=0}^{M-1} [k] \sum_{n=0}^{N-1} [l] h[m-k, n-l] x[k, l]$$

for m = 0, ..., M-1 and n = 0, ..., N-1. Note that  $-(M-1) \le m-k \le M-1$  and  $-(N-1) \le n-l \le N-1$ in the above convolution expression. We would like to express that convolution as a **circular convolution** so that it can be implemented using FFT operations. We embed both x and h in 2D arrays of size  $2M \times 2N$  as follows:

$$x_{2}[m,n] = \begin{cases} x[m,n], & m = 0, \dots, M-1, \ n = 0, \dots, N-1 \\ 0, & m = M, \dots, 2M-1 \text{ or } n = N, \dots, 2N-1, \end{cases}$$
$$h_{2}[m,n] = \begin{cases} h_{0}[m,n], & m = 0, \dots, 2M-1, \quad n = 0, \dots, 2N-1, \quad m \neq M, \quad n \neq N \\ 0, & m = M \text{ or } n = N, \end{cases}$$

where

$$h_0[m,n] \triangleq h[((m+M) \operatorname{mod} 2M) - M, ((n+N) \operatorname{mod} 2N) - N].$$

Now define the  $2M \times 2N$  point circular convolution of  $h_2$  and  $x_2$  as follows:

$$y_2[m,n] = h_2[m,n] \circledast_{(2M,2N)} x_2[m,n] = \sum_{k=0}^{2M-1} \sum_{l=0}^{2N-1} h_2[(m-k) \operatorname{mod} 2M, \ (n-l) \operatorname{mod} 2N] x_2[k,l].$$

Then one can verify that  $y[m,n] = y_2[m,n]$  for m = 0, ..., M - 1 and n = 0, ..., N - 1. So we can compute y[m,n] as follows.

• Precompute the  $2M \times 2N$  array  $h_2[m, n]$  and its 2D FFT  $H_2[k, l]$ .

• Zero pad the signal x[m, n] to size  $2M \times 2N$  and compute its 2D FFT X[k, l].

fig\_nufft\_emax\_core\_

• Multiply  $H_2$  and X to form  $Y_2[k, l] = H_2[k, l]X[k, l]$ . Compute the inverse 2D FFT of  $Y_2$  and extract the 1st  $M \times N$  quadrant.

For iterative image reconstruction, we must computing Tx several times. Using the above recipe, each such multiplication requires only one 2D FFT and one inverse 2D FFT (not counting the single 2D FFT that is needed to form  $H_2[k, l]$  prior to iterating). See also [194].

#### 6.9 **Problems** (s,four,prob)

s,four,prob

**Problem 6.1** For  $\mathcal{A}$  defined in (6.1.8), show that  $\|\mathcal{A}\|_2 \leq \sqrt{n_d} |\mathcal{D}| = \|\mathcal{A}\|_{\text{Frob}}$  where  $|\mathcal{D}| = \int_{\mathcal{D}} d\vec{x}$ . Find analogous results for A in (6.2.4).

**Problem 6.2** For the case of Cartesian sampled k-space data with under-sampling, show that the minimum-norm *least squares (MNLS)* solution (6.3.4) is equivalent to zero-filling the missing k-space data and applying an inverse FT or DFT.

**Problem 6.3** Variations on the following minimization problem arise repeatedly in §6.4.2:

 $\hat{\boldsymbol{w}} = \operatorname*{arg\,min}_{\boldsymbol{w}} \Psi(\boldsymbol{w}), \quad \Psi(\boldsymbol{w}) = \|\boldsymbol{C} - \boldsymbol{B} \, \boldsymbol{D}(\boldsymbol{w}) \, \boldsymbol{A} \|_{\mathrm{Frob}}^2.$ 

Using the Frobenius norm property (26.5.6) that  $\|A\|_{\text{Frob}}^2 = \text{trace}\{A'A\}$ , show that

$$\Psi(\boldsymbol{w}) = \operatorname{trace}\{\boldsymbol{C}'\boldsymbol{C}\} - 2\operatorname{real}\{\boldsymbol{w}'\boldsymbol{v}\} + \boldsymbol{w}'\boldsymbol{M}\boldsymbol{w}$$
(6.9.1)

$$M_{li} = [\mathbf{B}'\mathbf{B}]_{li}[\mathbf{A}\mathbf{A}']_{il}, \quad l, i = 1, ..., n_{d}$$
(6.9.1)
$$M_{li} = [\mathbf{B}'\mathbf{B}]_{li}[\mathbf{A}\mathbf{A}']_{il}, \quad l, i = 1, ..., n_{d}$$
(6.9.2)

$$v_i = [B'CA']_{ii}, \quad i = 1, \dots, n_d.$$
 (6.9.3)

Show furthermore that the Hermitian symmetric matrix M is positive-semidefinite, and that if M u = 0, then u' v =0, i.e.,  $\boldsymbol{v} \in \mathcal{R}_{\boldsymbol{M}}$ .

Show therefore that the minimizer of  $\Psi(w)$  satisfies Mw = v. *Note: if one wants to restrict the minimization to real-valued* w*, then the solution satisfies* real{M}  $w = real{v}$ .

**Problem 6.4** *Prove the gridding replication property* (6.4.57).

**Problem 6.5** Find choices for the weighting matrices  $W_1$  and  $W_2$  in the data-domain criterion (6.4.20) that lead to solutions corresponding to each of the examples of  $\hat{w}$  for the image domain criteria in §6.4.2.3.4. (In other words, these two formulations are closely related.)

Problem 6.6 For (Cartesian) partial k-space data as shown in Fig. 6.1.1, compare three types of reconstruction approaches: (i) one of the methods described in [195], (ii) an iterative regularized LS estimation approach based on the standard quadratic roughness penalty in (1.10.1), and (iii) based on a regularizer that penalizes the magnitude and phase separately [196]. Focus on the following model:

$$y = Af + \varepsilon$$
,

where  $f = D(\phi)x$  and  $D(\phi) = \text{diag}\{e^{i\phi_j}\}$ .

The standard approach computes a "low-resolution" phase estimate  $\phi$  from the central k-space data, which may be reasonable when the phase is expected to be spatially smooth. (Here, low resolution means truncated k-space, but still the number of pixels matches that of the image x.) Given this phase estimate, there are several approaches for reconstructing x.

*POCS* approach (ignores noise): find  $\hat{f} \in C_1 \cap C_2$  where

$$\mathcal{C}_1 riangleq \{oldsymbol{f} \in \mathbb{C}^{n_{ ext{p}}} : oldsymbol{A}oldsymbol{f} = oldsymbol{y}\}$$

and

$$\mathcal{C}_2 \triangleq \left\{ \boldsymbol{f} \in \mathbb{C}^{n_{\mathrm{p}}} : \angle f_j = \hat{\phi}_j, \ j = 1, \dots, n_{\mathrm{p}} \right\}.$$

Iterative approach based on **phase correction**:  $\hat{f} = D(\hat{\phi}) \hat{x}$ , where

$$\hat{\boldsymbol{x}} = \operatorname*{arg\,min}_{\boldsymbol{x} \in \mathbb{R}^{n_{\mathrm{p}}}} rac{1}{2} \left\| \boldsymbol{y} - \boldsymbol{A} \boldsymbol{D}(\hat{\boldsymbol{\phi}}) \boldsymbol{x} \right\|_{2}^{2} + \beta \operatorname{\mathsf{R}}(\boldsymbol{x})$$

Bydder approach [197]:  $\hat{f} = D(\hat{\phi}) \hat{x}$ , where

$$\underset{\boldsymbol{x} \in \mathbb{C}^{n_{\mathrm{p}}}}{\arg\min} \frac{1}{2} \left\| \boldsymbol{y} - \boldsymbol{A} \boldsymbol{D}(\hat{\boldsymbol{\phi}}) \boldsymbol{x} \right\|_{2}^{2} + \alpha \left\| \max\{\boldsymbol{x}\} \right\|^{2} + \beta \operatorname{\mathsf{R}}(\boldsymbol{x})$$

Joint estimation approach (based on phase regularization) [196]:

$$\underset{\boldsymbol{x} \in \mathbb{R}^{n_{\mathrm{p}}}, \boldsymbol{\phi} \in \mathbb{R}^{n_{\mathrm{p}}}}{\arg\min} \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{D}(\boldsymbol{\phi})\boldsymbol{x}\|_{2}^{2} + \alpha \, \mathsf{R}(\boldsymbol{\phi}) + \beta \, \mathsf{R}(\boldsymbol{x}),$$

where  $\alpha$  is "large" to provide a smooth phase estimate.

(Solve?)

**Problem 6.7** Prove that the functions  $\phi_i$  in (6.1.4) are linearly independent. Hint. One simple proof for the 1D case where  $\bar{d} = 1$  uses the fact that Vandermonde matrices are invertible. If needed, assume that  $\vec{0} \in D$ .

**Problem 6.8** Implement iterative reconstruction from Fourier samples using Dirac, rect, and sinc bases as described in §6.2.6 and compare reconstructed image quality. Use a true object that consists of shapes with analytical Fourier transform expressions such as triangles. (Solve?)

**Problem 6.9** Analyze  $\mathsf{E}\begin{bmatrix}\hat{f}\end{bmatrix}$  for the nonparametric QPWLS estimator in (6.3.7) for the case where W = I and the frequencies  $\{\vec{v}_i\}$  are equally spaced. (Solve?)

**Problem 6.10** Considering (6.2.9) and (6.2.10) as an approximation to (6.3.7), compute (and display) some of the  $g_i$  functions in (6.3.8). Compare the nature of these "basis functions" to the usual choices such as pixels and B-splines. (Solve?)

**Problem 6.11** Using (6.2.13), for the unweighted case with W = I, and no regularization ( $\mathbf{R} = \mathbf{0}$ ), the mean of the QPWLS estimator is  $\mathsf{E}[\hat{x}] = [\mathbf{E}'\mathbf{B}'\mathbf{B}\mathbf{E}]^{-1}\mathbf{E}'\mathbf{B}'\mathbf{A}f$ . Suppose the frequency samples are equally spaced and  $\mathbf{E}$  is orthonormal. If f is a rect function of width  $\Delta$ , compare  $\mathsf{E}[\hat{x}]$  for a model where  $\mathbf{B}$  corresponds to  $n_{\mathrm{p}} \gg 1$  rect basis functions of width  $\Delta$ , and a model where  $\mathbf{B}$  corresponds to Dirac impulse functions spaced  $\Delta$  apart.

, four, a, exp Pro

p,four,qpwls,B

four, dens, sind

**Problem 6.12** For the complex exponential basis (6.2.15), determine the general form of the elements of the system matrix  $\mathbf{A}$  in (6.2.4). Then consider the specific 1D case where  $\mathcal{D} = [-\text{FOV}/2, \text{FOV}/2]$  and  $n_d = n_p$ , and the frequency samples  $\vec{\nu}_i$  are uniformly spaced with  $\Delta_{\nu} = 1/\text{FOV}$ . Comment on the relative complexity of implementing iterative reconstruction from nonuniform frequency samples using a complex exponential basis versus using equally spaced basis functions such as rects.

**Problem 6.13** *Prove the conjecture about diminishing basis function widths at the end of* §6.2.11. (Solve?)

**Problem 6.14** Requiring  $\mathcal{W}$  in (6.3.1) to map into  $\mathcal{L}_2(\mathcal{D})$  is a subtle restriction because it would seem to preclude "convolution-like" characteristics that might expand the support of a function originally in  $\mathcal{L}_2(\mathcal{D})$ . We might prefer to assume only  $\mathcal{W} : \mathcal{L}_2 \to \mathcal{L}_2$ , but this would appear to require allowing  $\hat{f} \in \mathcal{L}_2(\mathbb{R}^{\bar{d}})$ , thereby losing the benefit of the support constraint. Resolve this dilemma. (Solve?)

**Problem 6.15** *Prove the sinc property* (6.4.8).

**Problem 6.16** Prove the simplified image domain criterion (6.4.19).

**Problem 6.17** Prove the "generalized Sedarat" solution (6.4.24). (It may be easier to use a discrete approximation A to A.)

**Problem 6.18** *Prove that* (6.4.24) *is more general than* (6.4.21) *even in the special case where we choose*  $\mathcal{Z} = \mathcal{K}\mathcal{A}^*$ .

**Problem 6.19** *Prove the "weighted Sedarat" solution* (6.4.28).

(Need typed.)

Problem 6.20 Prove the regularized solution (6.4.43).

**Problem 6.21** Consider a (2D) polar sampling pattern in Fourier space with samples that are equally spaced by  $\Delta_{\nu}$ along  $\rho$  with  $N_{\rm A}$  equally spaced radial lines over  $2\pi$ . Determine analytically some reasonable values for density compensation factors for samples where  $\rho = k \Delta_{\nu}$ , for k = 0, 1, ... Hint: think of (approximate) Voronoi cells for  $N_{\rm A}$  large. See [198–200]. **Problem 6.22** Given samples of a function f(x, y) on equally spaced rings  $\{f(n\Delta, \phi) : n \in \mathbb{Z}, \phi \in [0, 2\pi)\}$ , determine conditions under which f(x, y) can be recovered and describe a recovery procedure. (This problem is related to the PSF of spiral and radial sampling; see [198].)

(Solve?)

**Problem 6.23** Determine analytically some reasonable values for density compensation factors for the linogram sampling pattern shown in Fig. 3.8.1. (Solve?)

**Problem 6.24** Using ideas similar to \$25.7.12, develop an analytical approximation to the condition number of A'WA for the case of radial samping.

(Solve?)

**Problem 6.25** Use the fast summation method in [201] to develop a fast (?) gridding algorithm of the form (6.4.51) or (6.4.52) even for a kernel with a large support. (Solve?)

**Problem 6.26** Derive the ideal real-valued interpolators (6.6.5) and (6.6.10). Optional: also attempt a similar derivation when s[N - 1] = s[n].

## 6.10 Bibliography

hogbom:74:asw

vray:01:sab

- [1] R. N. Bracewell and A. R. Thompson. "The main beam and ring lobes of an east-west rotation-synthesis array." In: *The Astrophysical Journal* 182 (May 1973), 77–94. URL: http://adsabs.harvard.edu/cgi-bin/nph-bib\_query?bibcode=1973ApJ...182...77B (cit. on pp. 6.2, 6.18).
- [2] J. A. Högbom. "Aperture synthesis with a non-regular distribution of interferometer baselines." In: Astronomy and Astrophysics Supplement 15 (1974), 417–26. URL: http://adsabs.harvard.edu/cgibin/nph-bib\_query?bibcode=1974A&AS...15..417H (cit. on p. 6.2).
- [3] U. J. Schwarz. "Mathematical-statistical description of the iterative beam removing technique (Method CLEAN)." In: Astronomy and Astrophysics 65 (1978), 345–56. URL: http: //adsabs.harvard.edu/cgi-bin/nph-bib\_query?bibcode=1978A&A....65...345S (cit. on p. 6.2).
- [4] J. J. Clark, M. R. Palmer, and P. D. Lawrence. "A transformation method for the reconstruction of functions from nonuniformly spaced samples." In: *IEEE Trans. Acoust. Sp. Sig. Proc.* 33.5 (Oct. 1985), 1151–65. DOI: 10.1109/TASSP.1985.1164714 (cit. on pp. 6.2, 6.28).
- [5] A. R. Thompson, J. M. Moran, and G. W. Swenson. *Interferometry and synthesis in radio astronomy*. 3rd edition. Springer, 2017. DOI: 10.1007/978-3-319-44431-4 (cit. on p. 6.2).
- [6] D. C. Munson, J. D. O'Brien, and W. K. Jenkins. "A tomographic formulation of spotlight mode synthetic aperture radar." In: *Proc. IEEE* 71.8 (Aug. 1983), 917–25. DOI: 10.1109/PROC.1983.12698 (cit. on p. 6.2).
  - [7] D. C. Munson and J. L. Sanz. "Image reconstruction from frequency-offset Fourier data." In: *Proc. IEEE* 72.6 (June 1984), 661–9. DOI: 10.1109/PROC.1984.12915 (cit. on p. 6.2).
- [8] Y. Wu and D. Munson. "Multistatic passive radar imaging using the smoothed pseudo Wigner-Ville distribution." In: *Proc. IEEE Intl. Conf. on Image Processing*. Vol. 3. 2001, 604–7. DOI: 10.1109/ICIP.2001.958191 (cit. on p. 6.2).
- [9] T. J. Kragh and A. A. Kharbouch. "Monotonic iterative algorithms for SAR image restoration." In: *Proc. IEEE Intl. Conf. on Image Processing*. 2006, 645–8. DOI: 10.1109/ICIP.2006.312463 (cit. on p. 6.2).
- [10] S. Sun, G. Zhu, and T. Jin. "Novel methods to accelerate CS radar imaging by NUFFT." In: *IEEE Trans. Geosci. Remote Sensing* 53.1 (Jan. 2015), 557–66. DOI: 10.1109/TGRS.2014.2325492 (cit. on p. 6.2).
  - [11] L. Feng et al. "The gridding method for image reconstruction of nonuniform aperture synthesis radiometers." In: *IEEE Geoscience and Remote Sensing Letters* 12.2 (Feb. 2015), 274–78. DOI: 10.1109/LGRS.2014.2335413 (cit. on p. 6.2).
  - [12] D. Vray et al. "Synthetic aperture-based beam compression for intravascular ultrasound imaging." In: *IEEE Trans. Ultrasonics, Ferro., and Freq. Control* 48.1 (Jan. 2001), 189–201. DOI: 10.1109/58.896132 (cit. on p. 6.2).

haun:02:etd	[13]	M. A. Haun, D. L. Jones, and W. D. O'Brien. "Efficient three-dimensional imaging from a small cylindrical aperture." In: <i>IEEE Trans. Ultrasonics, Ferro., and Freq. Control</i> 49.7 (July 2002), 861–70. DOI: 10.1109/TUFFC.2002.1020156 (cit. on p. 6.2).
bronstein:U2:ria	[14]	M. M. Bronstein et al. "Reconstruction in diffraction ultrasound tomography using nonuniform FFT." In: <i>IEEE Trans. Med. Imag.</i> 21.11 (Nov. 2002), 1395–1401. DOI: 10.1109/TMI.2002.806423 (cit. on p. 6.2).
twieg:83:tkt	[15]	D. B. Twieg. "The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods." In: <i>Med. Phys.</i> 10.5 (Sept. 1983), 610–21. DOI: 10.1118/1.595331 (cit. on p. 6.2).
jackson:91:soa	[16]	J. I. Jackson et al. "Selection of a convolution function for Fourier inversion using gridding." In: <i>IEEE Trans.</i> <i>Med. Imag.</i> 10.3 (Sept. 1991), 473–8. DOI: 10.1109/42.97598 (cit. on pp. 6.2, 6.19, 6.21, 6.31).
pipe:99:sdc	[17]	J. G. Pipe and P. Menon. "Sampling density compensation in MRI: Rationale and an iterative numerical solution." In: <i>Mag. Res. Med.</i> 41.1 (Jan. 1999), 179–86. DOI: 10.1002/(SICI)1522-2594(199901)41:1<179::AID-MRM25>3.0.CO; 2-V (cit. on pp. 6.2, 6.24, 6.25)
sutton:03:fii	[18]	B. P. Sutton, D. C. Noll, and J. A. Fessler. "Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities." In: <i>IEEE Trans. Med. Imag.</i> 22.2 (Feb. 2003), 178–88. DOI: 10.1109/TMI.2002.808360 (cit. on pp. 6.2, 6.10).
press:89:faf	[19]	W. H. Press and G. B. Rybicki. "Fast algorithm for spectral analysis of unevenly sampled data." In: <i>The Astrophysical Journal</i> 338 (Mar. 1989), 227–80. DOI: 10.1086/167193 (cit. on p. 6.2).
larsson:01:sic	[20]	E. Larsson, P. Stoica, and J. Li. "SAR image construction from gapped phase-history data." In: <i>Proc. IEEE</i> <i>Intl. Conf. on Image Processing</i> . Vol. 3. 2001, 608–11. DOI: 10.1109/ICIP.2001.958192 (cit. on p. 6.2).
yudilevich:87:ifs	[21]	<ul> <li>E. Yudilevich and H. Stark. "Interpolation from samples on a linear spiral scan." In: <i>IEEE Trans. Med. Imag.</i></li> <li>6.3 (Sept. 1987), 193–200. DOI: 10.1109/TMI.1987.4307827 (cit. on p. 6.2).</li> </ul>
yudilevich:88:sst	[22]	E. Yudilevich and H. Stark. "Spiral sampling: theory and an application to magnetic resonance imaging." In: <i>J. Opt. Soc. Am. A</i> 5.4 (Apr. 1988), 542–53. DOI: 10.1364/JOSAA.5.000542 (cit. on p. 6.2).
soumekh:89:ras	[23]	M. Soumekh. "Reconstruction and sampling constraints for spiral data." In: <i>IEEE Trans. Acoust. Sp. Sig. Proc.</i> 37.6 (June 1989), 882–91. DOI: 10.1109/ASSP.1989.28059 (cit. on pp. 6.2, 6.18).
myridis:98:soc	[24]	N. E. Myridis and C. Chamzas. "Sampling on concentric circles." In: <i>IEEE Trans. Med. Imag.</i> 17.2 (Apr. 1998), 294–9. DOI: 10.1109/42.700741 (cit. on p. 6.2).
edholm:88:irf	[25]	P. Edholm, G. T. Herman, and D. A. Roberts. "Image reconstruction from linograms: implementation and evaluation." In: <i>IEEE Trans. Med. Imag.</i> 7.3 (Sept. 1988), 239–46. DOI: 10.1109/42.7788 (cit. on p. 6.2).
lawton:88:anp	[26]	W. Lawton. "A new polar Fourier transform for computer-aided tomography and spotlight synthetic aperture radar." In: <i>IEEE Trans. Acoust. Sp. Sig. Proc.</i> 36.6 (June 1988), 931–3. DOI: 10.1109/29.1609 (cit. on p. 6.2).
herman:92:ftd	[27]	G. T. Herman, D. Roberts, and L. Axel. "Fully three-dimensional reconstruction from data collected on concentric cubes in Fourier space: implementation and a sample application to MRI." In: <i>Phys. Med. Biol.</i> 37.3 (Mar. 1992), 673–89. DOI: 10.1088/0031-9155/37/3/013 (cit. on p. 6.2).
dusaussoy:96:vav	[28]	N. J. Dusaussoy. "VOIR: a volumetric image reconstruction algorithm based on Fourier techniques for inversion of the 3-D Radon transform." In: <i>IEEE Trans. Im. Proc.</i> 5.1 (Jan. 1996), 121–31. DOI: 10.1109/83.481676 (cit. on p. 6.2).
bellon:97:fdf	[29]	P. L. Bellon and S. Lanzavecchia. "Fast direct Fourier methods, based on one- and two-pass coordinate transformations, yield accurate reconstructions of x-ray CT clinical images." In: <i>Phys. Med. Biol.</i> 42.3 (Mar. 1997), 443–64. DOI: 10.1088/0031-9155/42/3/001 (cit. on p. 6.2).
seger:98:tdr potts:01:anl	[30]	M. M. Seger. "Three-dimensional reconstruction from cone-beam data using an efficient Fourier technique combined with a special interpolation filter." In: <i>Phys. Med. Biol.</i> 43.4 (Apr. 1998), 951–60. DOI: 10.1088/0031-9155/43/4/022 (cit. on p. 6.2).

[31] D. Potts and G. Steidl. "A new linogram algorithm for computerized tomography." In: *IMA J. Numer. Anal.* 21.3 (July 2001), 769–82. DOI: 10.1093/imanum/21.3.769 (cit. on p. 6.2).

averbuch:08:aff-1	[32]	A. Averbuch et al. "A framework for discrete integral transformations I—The pseudopolar Fourier transform." In: <i>SIAM J. Sci. Comp.</i> 30.2 (2008), 764–84. DOI: 10.1137/060650283 (cit. on p. 6.2).
averbuch:08:aff-2	[33]	A. Averbuch et al. "A framework for discrete integral transformations II—The 2D discrete Radon transform." In: <i>SIAM J. Sci. Comp.</i> 30.2 (2008), 785–803. DOI: 10.1137/060650301 (cit. on p. 6.2).
averbuch:02:tpf	[34]	A. Averbuch et al. <i>The pseudopolar FFT and its applications</i> . 2002. URL:
liang:92:crm		http://www.cs.tau.ac.il/~amirl/publications.html (cit. on p. 6.2).
li:15:aln	[35]	Z. P. Liang et al. "Constrained reconstruction methods in MR imaging." In: <i>Reviews of Magnetic Resonance in Medicine</i> 4 (1992), 67–185 (cit. on p. 6.3).
	[36]	G. Li et al. "An L1-norm phase constraint for half-Fourier compressed sensing in 3D MR imaging." In: <i>Mag. Res. Mat. Bio. Phys. Med.</i> 28.5 (Oct. 2015), 459–72. DOI: 10.1007/s10334-015-0482-7 (cit. on p. 6.3).
lewitt:03:00m	[37]	R. M. Lewitt and S. Matej. "Overview of methods for image reconstruction from projections in emission computed tomography." In: <i>Proc. IEEE</i> 91.10 (Oct. 2003), 1588–611. DOI: 10.1109/JPROC.2003.817882 (cit. on p. 6.3).
haacKe:8/:sin	[38]	E. M. Haacke. "Solving for non-ideal conditions in two-dimensional Fourier transform magnetic resonance imaging using a generalised inverse transform." In: <i>Inverse Prob.</i> 3.3 (Aug. 1987), 421–35. DOI: 10.1088/0266-5611/3/3/011 (cit. on p. 6.4).
duffin:52:aco	[39]	R. J. Duffin and A. C. Schaeffer. "A class of nonharmonic Fourier series." In: <i>Trans. Amer. Math. Soc.</i> 72.2 (Mar. 1952), 341–66. URL: http://www.jstor.org/stable/1990760 (cit. on p. 6.5).
grochenig:93:aot	[40]	K. Gröchenig. "Acceleration of the frame algorithm." In: <i>IEEE Trans. Sig. Proc.</i> 41.12 (Dec. 1993), 3331–40. DOI: 10.1109/78.258077 (cit. on p. 6.5).
feichtinger:95:enm	[41]	H. G. Feichtinger, K. Gröchenig, and T. Strohmer. "Efficient numerical methods in non-uniform sampling theory." In: <i>Numerische Mathematik</i> 69.4 (Feb. 1995), 423–40. DOI: 10.1007/s002110050101 (cit. on pp. 6.5, 6.7, 6.10, 6.30).
barrett:95:oao	[42]	H. H. Barrett et al. "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance." In: <i>J. Opt. Soc. Am. A</i> 12.5 (May 1995), 834–52. DOI: 10.1364/JOSAA.12.000834 (cit. on p. 6.5).
luenberger:69	[43]	D. G. Luenberger. <i>Optimization by vector space methods</i> . New York: Wiley, 1969. URL: http://books.google.com/books?id=1ZU0CAH4RccC (cit. on pp. 6.6, 6.13, 6.15).
grochenig:Ul:nat	[44]	K. Gröchenig and T. Strohmer. "Numerical and theoretical aspects of non-uniform sampling of band-limited images." In: <i>Nonuniform Sampling: Theory and Practice</i> . Ed. by F. Marvasti. New York: Kluwer, 2001, pp. 283–324 (cit. on pp. 6.7, 6.10, 6.11, 6.18, 6.28).
grochenig:01:nus	[45]	K. Gröchenig. "Non-uniform sampling in higher dimensions: from trigonometric polynomials to band-limited functions." In: <i>Modern Sampling Theory: Mathematics and Applications</i> . Ed. by J J Benedetto and P Ferreira. Boston: Birkhauser, 2001, pp. 155–71 (cit. on p. 6.7).
sidiropoulos:Ul:gcu	[46]	N. D. Sidiropoulos. "Generalizing Carathéodory's uniqueness of harmonic parameterization to $N$ dimensions." In: <i>IEEE Trans. Info. Theory</i> 47.4 (May 2001), 1687–90. DOI: 10.1109/18.923759 (cit. on p. 6.7).
bass:04:rso	[47]	R. F. Bass and K. Gröchenig. "Random sampling of multivariate trigonometric polynomials." In: <i>SIAM J. Math. Anal.</i> 36.3 (2004), 773–95. DOI: 10.1137/S0036141003432316 (cit. on p. 6.7).
fessler:96:srp	[48]	J. A. Fessler and W. L. Rogers. "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs." In: <i>IEEE Trans. Im. Proc.</i> 5.9 (Sept. 1996), 1346–58. DOI: 10.1109/83.535846 (cit. on p. 6.8).
dharanipragada:96:rli	[49]	S. Dharanipragada and K. S. Arun. "Resolution limits in signal recovery." In: <i>IEEE Trans. Sig. Proc.</i> 44.3 (Mar. 1996), 546–61. DOI: 10.1109/78.489028 (cit. on p. 6.8).
fessler:03:nff	[50]	J. A. Fessler and B. P. Sutton. "Nonuniform fast Fourier transforms using min-max interpolation." In: <i>IEEE Trans. Sig. Proc.</i> 51.2 (Feb. 2003), 560–74. DOI: 10.1109/TSP.2002.807005 (cit. on pp. 6.10, 6.27, 6.35, 6.36, 6.37, 6.38, 6.40, 6.41).
Dourgeols:01:rom	[51]	M. Bourgeois et al. "Reconstruction of MRI images from non-uniform sampling and application to intrascan motion correction in functional MRI." In: <i>Modern Sampling Theory: Mathematics and Applications</i> . Ed. by J J Benedetto and P Ferreira. Boston: Birkhauser, 2001, pp. 343–63 (cit. on p. 6.10).

feron:03:bef

strang:86:apf

wax:83:eio

wang:99:aim

- [52] K. P. Pruessmann et al. "Advances in sensitivity encoding with arbitrary k-space trajectories." In: *Mag. Res. Med.* 46.4 (Oct. 2001), 638–51. DOI: 10.1002/mrm.1241 (cit. on p. 6.10).
  - [53] B. Desplanques et al. "Iterative reconstruction of magnetic resonance images from arbitrary samples in k-space." In: *IEEE Trans. Nuc. Sci.* 49.5-1 (Oct. 2002), 2268–73. DOI: 10.1109/TNS.2002.803673 (cit. on p. 6.10).
  - [54] L. Greengard and J-Y. Lee. "Accelerating the nonuniform fast Fourier transform." In: *SIAM Review* 46.3 (Sept. 2004), 443–54. DOI: 10.1137/S003614450343200X (cit. on pp. 6.10, 6.17, 6.38).
  - [55] F. Feron and K. Sauer. "Bayesian estimation for rheological MRI." In: *Proc. SPIE 5016 Computational Imaging*. 2003, 123–30. DOI: 10.1117/12.479703 (cit. on p. 6.10).
  - [56] Y. M. Kadah et al. "Progressive magnetic resonance image reconstruction based on iterative solution of a sparse linear system." In: *Intl. J. Biomedical Im.* 2006 (2006), p. 49378. DOI: 10.1155/IJBI/2006/49378 (cit. on p. 6.10).
  - [57] G. Strang. "A proposal for Toeplitz matrix calculations." In: Stud. Appl. Math. 74 (1986), 171–6 (cit. on p. 6.10).
  - [58] R. H. Chan and M. K. Ng. "Conjugate gradient methods for Toeplitz systems." In: *SIAM Review* 38.3 (Sept. 1996), 427–82. DOI: 10.1137/S0036144594276474 (cit. on pp. 6.10, 6.11).
    - [59] T. Strohmer. "Computationally attractive reconstruction of bandlimited images from irregular samples." In: *IEEE Trans. Im. Proc.* 6.4 (Apr. 1997), 540–8. DOI: 10.1109/83.563319 (cit. on pp. 6.10, 6.28).
    - [60] A. J. W. Duijndam, M. A. Schonewille, and C. O. H. Hindriks. "Reconstruction of band-limited signals, irregularly sampled along one spatial direction." In: *Geophysics* 64.2 (Mar. 1999), 524–38. DOI: 10.1190/1.1444559 (cit. on p. 6.10).
    - [61] T. Strohmer. "Numerical analysis of the non-uniform sampling problem." In: J. Comp. Appl. Math. 122.1-2 (Oct. 2000), 297–316. DOI: 10.1016/S0377-0427 (00) 00361-7 (cit. on pp. 6.10, 6.28).
    - [62] F. Wajer and K. P. Pruessmann. "Major speedup of reconstruction for sensitivity encoding with arbitrary trajectories." In: *Proc. Intl. Soc. Mag. Res. Med.* 2001, p. 767. URL: http://cds.ismrm.org/ismrm-2001/PDF3/0767.pdf (cit. on p. 6.11).
      - [63] H. Eggers, P. Boernert, and P. Boesiger. "Comparison of gridding- and convolution-based iterative reconstruction algorithms for sensitivity-encoded non-cartesian acquisitions." In: *Proc. Intl. Soc. Mag. Res. Med.* 2002, p. 743. URL: http://cds.ismrm.org/ismrm-2002/PDF3/0743.PDF (cit. on p. 6.11).
      - [64] C. Liu, M. E. Moseley, and R. Bammer. "Fast SENSE reconstruction using linear system transfer function." In: Proc. Intl. Soc. Mag. Res. Med. 2005, p. 689. URL: http://cds.ismrm.org/ismrm-2005/Files/00689.pdf (cit. on p. 6.11).
      - [65] C. Baron et al. "Rapid non-cartesian regularized SENSE reconstruction using a point spread function model." In: *Proc. Intl. Soc. Mag. Res. Med.* 2017, p. 4002 (cit. on p. 6.11).
      - [66] M. Wax and T. Kailath. "Efficient inversion of Toeplitz-block Toeplitz matrix." In: *IEEE Trans. Acoust. Sp. Sig. Proc.* 31.5 (Oct. 1983), 1218–21. DOI: 10.1109/TASSP.1983.1164208 (cit. on p. 6.11).
        - [67] E. Linzer and M. Vetterli. "Structure and arithmetic complexity of products and inverses of Toeplitz matrices." In: *Linear Algebra and its Applications* 145 (Feb. 1991), 159–75. DOI: 10.1016/0024-3795 (91) 90293-6 (cit. on p. 6.11).
      - [68] S. M. Kay. *Fundamentals of statistical signal processing: Estimation theory*. New York: Prentice-Hall, 1993 (cit. on pp. 6.11, 6.15).
      - [69] K. P. Pruessmann et al. "SENSE: sensitivity encoding for fast MRI." In: *Mag. Res. Med.* 42.5 (Nov. 1999), 952–62. DOI: '10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S' (cit. on p. 6.13).
      - [70] S. J. Wernecke and L. R. DAddario. "Maximum entropy image reconstruction." In: *IEEE Trans. Comp.* 26.4 (Apr. 1977), 351–64. DOI: 10.1109/TC.1977.1674845 (cit. on p. 6.13).
      - [71] J. C. Hoch et al. "Maximum entropy reconstruction of complex (phase-sensitive) spectra." In: *J. Mag. Res.* 86.2 (Feb. 1990), 236–46. DOI: 10.1016/0022-2364 (90) 90256-9 (cit. on p. 6.13).
      - [72] Y. Wang and X. Zhao. "An iterative method for maximum entropy regularization reconstruction in MRI." In: *Intl. J. Imaging Sys. and Tech.* 10.6 (Mar. 1999), 427–31. DOI:
         ' 10.1002/(SICI)1098-1098(1999)10:6<427::AID-IMA3>3.0.CO; 2-7' (cit. on p. 6.13).

bertero:85:lip

hsieh:96:iru

- [73] R. T. Constable and R. M. Henkelman. "Why MEM does not work in MR image reconstruction." In: *Mag. Res. Med.* 14.1 (Apr. 1990), 12–25. DOI: 10.1002/mrm.1910140103 (cit. on p. 6.13).
- [74] A. Manduca et al. "Autocorrection in MR imaging: Adaptive motion correction without navigator echoes." In: *Radiology* 215.3 (June 2000), 904–9. DOI: 10.1148/radiology.215.3.r00jn19904 (cit. on p. 6.13).
- [75] A. C. Gilbert, S. Muthukrishnan, and M. J. Strauss. "Improved time bounds for near-optimal sparse Fourier representations." In: *Proc. SPIE 5914 Wavelets XI*. 2005, 59141A. DOI: 10.1117/12.615931 (cit. on p. 6.13).
- [76] A. C. Gilbert et al. "Sublinear approximation of signals." In: *Proc. SPIE 6232 Intelligent Integrated Microsystems*. 2006, p. 623206. DOI: 10.1117/12.669596 (cit. on p. 6.13).
  - [77] E. J. Candes, J. Romberg, and T. Tao. "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information." In: *IEEE Trans. Info. Theory* 52.2 (Feb. 2006), 489–509. DOI: 10.1109/TIT.2005.862083 (cit. on p. 6.13).
- [78] E. J. Candes, J. Romberg, and T. Tao. "Stable signal recovery from incomplete and inaccurate measurements." In: *Comm. Pure Appl. Math.* 59.8 (2006), 1207–23. DOI: 10.1002/cpa.20124 (cit. on p. 6.13).
- [79] M. Lustig et al. "Faster imaging with randomly perturbed, undersampled spirals and |L|1 reconstruction." In: Proc. Intl. Soc. Mag. Res. Med. 2005, p. 685. URL: http://cds.ismrm.org/ismrm-2005/Files/00685.pdf (cit. on p. 6.13).
- [80] M. Lustig et al. "k-t SPARSE: high frame rate dynamic MRI exploiting spatio-temporal sparsity." In: Proc. Intl. Soc. Mag. Res. Med. 2006, p. 2420. URL: http://dev.ismrm.org//2006/2420.html (cit. on p. 6.13).
- [81] M. Lustig, D. L. Donoho, and J. M. Pauly. "Rapid MR imaging with "compressed sensing" and randomly under-sampled 3DFT trajectories." In: *Proc. Intl. Soc. Mag. Res. Med.* 2006, p. 695. URL: http://dev.ismrm.org//2006/0695.html (cit. on p. 6.13).
- [82] V. M. Patel et al. "Gradient-based image recovery methods from incomplete Fourier measurements." In: IEEE Trans. Im. Proc. 21.1 (Jan. 2012), 94–105. DOI: 10.1109/TIP.2011.2159803 (cit. on p. 6.13).
- [83] S. Bonettini and M. Prato. "Nonnegative image reconstruction from sparse Fourier data: A new deconvolution algorithm." In: *Inverse Prob.* 26.9 (Sept. 2010), p. 095001. DOI: 10.1088/0266-5611/26/9/095001 (cit. on p. 6.13).
- [84] M. Bertero, C. De Mol, and E. R. Pike. "Linear inverse problems with discrete data, I: General formulation and singular system analysis." In: *Inverse Prob.* 1.4 (Nov. 1985), 301–30. DOI: 10.1088/0266-5611/1/4/004 (cit. on p. 6.13).
- [85] R. Van de Walle et al. "Reconstruction of MR images from data acquired on a general non-regular grid by pseudoinverse calculation." In: *IEEE Trans. Med. Imag.* 19.12 (Dec. 2000), 1160–7. DOI: 10.1109/42.897806 (cit. on p. 6.14).
- [86] M. H. Buonocore, M. R. Brody, and A. Macovski. "Fast minimum variance estimator for limited angle CT image reconstruction." In: *Med. Phys.* 8.5 (Sept. 1981), 695–702. DOI: 10.1118/1.594838 (cit. on p. 6.14).
- [87] M. H. Buonocore, W. R. Brody, and A. Macovski. "A natural pixel decomposition for two-dimensional image reconstruction." In: *IEEE Trans. Biomed. Engin.* 28.2 (Feb. 1981), 69–78. DOI: 10.1109/TBME.1981.324781 (cit. on p. 6.14).
- [88] Y. L. Hsieh et al. "Image reconstruction using a generalized natural pixel basis." In: *IEEE Trans. Nuc. Sci.* 43.4 (Aug. 1996), 2306–19. DOI: 10.1109/23.531896 (cit. on p. 6.14).
- [89] J. C. Ye, Y. Bresler, and P. Moulin. "A self-referencing level-set method for image reconstruction from sparse Fourier samples." In: *Intl. J. Comp. Vision* 50.3 (Dec. 2002), 253–70. DOI: 10.1023/A:1020822324006 (cit. on p. 6.15).
- [90] V. Solo. "Exact Tikhonov regularisation for the limited data computed tomography problem." In: *Proc. IEEE Conf. Acoust. Speech Sig. Proc.* Vol. 4. 1995, 2289–90. DOI: 10.1109/ICASSP.1995.479948 (cit. on p. 6.15).

dale:01:arl

duijndam:99:nff

- [91] V. Solo. "Regularisation of the limited data computed tomography problem via the boundary element method." In: *Proc. IEEE Intl. Conf. on Image Processing*. Vol. 2. 1995, 430–2. DOI: 10.1109/ICIP.1995.537507 (cit. on p. 6.15).
- [92] B. W. Silverman. "Spline smoothing: the equivalent variable kernel method." In: *Ann. Stat.* 12.3 (1984), 898–916. DOI: 10.1214/aos/1176346710 (cit. on p. 6.15).
- [93] A. Macovski. "Volumetric NMR imaging with time-varying gradients." In: *Mag. Res. Med.* 2.1 (Feb. 1985), 29–40. DOI: 10.1002/mrm.1910020105 (cit. on p. 6.16).
- [94] D-Q. Chen, R. B. Marr, and P. C. Lauterbur. "Reconstruction from NMR data acquired with imaging gradients having arbitrary time dependence." In: *IEEE Trans. Med. Imag.* 5.3 (Sept. 1986), 162–4. DOI: 10.1109/TMI.1986.4307765 (cit. on pp. 6.16, 6.21).
  - [95] A. Maeda, K. Sano, and T. Yokoyama. "Reconstruction by weighted correlation for MRI with time-varying gradients." In: *IEEE Trans. Med. Imag.* 7.1 (Mar. 1988), p. 26. DOI: 10.1109/42.3926 (cit. on p. 6.16).
- [96] B. Dale, M. Wendt, and J. L. Duerk. "A rapid look-up table method for reconstructing MR images from arbitrary K-space trajectories." In: *IEEE Trans. Med. Imag.* 20.3 (Mar. 2001), 207–17. DOI: 10.1109/42.918471 (cit. on pp. 6.17, 6.31).
- [97] Y. Qian, J. Lin, and D. Jin. "Direct reconstruction of MR images from data acquired on a non-Cartesian grid using an equal-phase-line algorithm." In: *Mag. Res. Med.* 47.6 (June 2002), 1228–33. DOI: 10.1002/mrm.10165 (cit. on p. 6.17).
- [98] A. Dutt and V. Rokhlin. "Fast Fourier transforms for nonequispaced data." In: *SIAM J. Sci. Comp.* 14.6 (Nov. 1993), 1368–93. DOI: 10.1137/0914081 (cit. on pp. 6.17, 6.41).
  - [99] G. Beylkin. "On the fast Fourier transform of functions with singularities." In: *Applied and Computational Harmonic Analysis* 2.4 (Oct. 1995), 363–81. DOI: 10.1006/acha.1995.1026 (cit. on p. 6.17).
  - [100] A. Dutt and V. Rokhlin. "Fast Fourier transforms for nonequispaced data, II." In: *Applied and Computational Harmonic Analysis* 2.1 (Jan. 1995), 85–100. DOI: 10.1006/acha.1995.1007 (cit. on p. 6.17).
  - [101] C. Anderson and M. D. Dahleh. "Rapid computation of the discrete Fourier transform." In: *SIAM J. Sci. Comp.* 17.4 (July 1996), 913–9. DOI: 10.1137/0917059 (cit. on p. 6.17).
  - [102] G. Steidl. "A note on the fast Fourier transforms for nonequispaced grids." In: *Adv. in Comp. Math.* 9.3 (Nov. 1998), 337–52. DOI: 10.1023/A:1018901926283 (cit. on p. 6.17).
    - [103] A. F. Ware. "Fast approximate Fourier transforms for irregularly spaced data." In: *SIAM Review* 40.4 (Dec. 1998), 838–56. DOI: 10.1137/S003614459731533X (cit. on p. 6.17).
    - [104] A. J. W. Duijndam and M. A. Schonewille. "Nonuniform fast Fourier transform." In: *Geophysics* 64.2 (Mar. 1999), 539–51. DOI: 10.1190/1.1444560 (cit. on p. 6.17).
      - [105] N. Nguyen and Q. H. Liu. "The regular Fourier matrices and nonuniform fast Fourier transforms." In: SIAM J. Sci. Comp. 21.1 (1999), 283–93. DOI: 10.1137/S1064827597325712 (cit. on pp. 6.17, 6.38, 6.41).
        - [106] D. Potts, G. Steidl, and M. Tasche. "Fast Fourier transforms for nonequispaced data: A tutorial." In: *Modern Sampling Theory: Mathematics and Applications*. Ed. by J J Benedetto and P Ferreira. Boston: Birkhauser, 2001, pp. 253–74. URL:
        - http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.30.521 (cit. on p. 6.17).
        - [107] D. Ruiz-Antolin and A. Townsend. A nonuniform fast Fourier transform based on low rank approximation. 2017. URL: http://arxiv.org/abs/1701.04492 (cit. on p. 6.17).
        - [108] S. J. Norton. "Fast magnetic resonance imaging with simultaneously oscillating and rotating field gradients." In: *IEEE Trans. Med. Imag.* 6.1 (Mar. 1987), 21–31. DOI: 10.1109/TMI.1987.4307794 (cit. on p. 6.18).
- [109] M. Soumekh. "Band-limited interpolation from unevenly spaced sampled data." In: *IEEE Trans. Acoust. Sp. Sig. Proc.* 36.1 (Jan. 1988), 118–22. DOI: 10.1109/29.1497 (cit. on pp. 6.18, 6.28).
  - [110] C. J. Hardy, H. E. Cline, and P. A. Bottomley. "Correcting for nonuniform k-space sampling in two-dimensional NMR selective excitation." In: *J. Mag. Res.* 87.3 (May 1990), 639–45. DOI: 10.1016/0022-2364 (90) 90323-2 (cit. on p. 6.18).
- [111] C. H. Meyer et al. "Fast spiral coronary artery imaging." In: *Mag. Res. Med.* 28.2 (Dec. 1992), 202–13. DOI: 10.1002/mrm.1910280204 (cit. on pp. 6.18, 6.32).

spielman:95:mrf

bogo + 0.7 + dof	[112]	D. M. Spielman, J. M. Pauly, and C. H. Meyer. "Magnetic resonance fluoroscopy using spirals with variable sampling densities." In: <i>Mag. Res. Med.</i> 34.3 (Sept. 1995), 388–94. DOI: 10.1002/mrm.1910340316 (cit. on p. 6.18).
noge:97:dCl	[113]	R. D. Hoge, R. K. S. Kwan, and G. B. Pike. "Density compensation functions for spiral MRI." In: <i>Mag. Res. Med.</i> 38.1 (July 1997), 117–28. DOI: 10.1002/mrm.1910380117 (cit. on p. 6.18).
Sarty.37.inp	[114]	G. E. Sarty. "Natural k-plane coordinate reconstruction method for magnetic resonance imaging: mathematical foundations." In: <i>Intl. J. Imaging Sys. and Tech.</i> 8.6 (1997), 519–28. DOI: '10.1002/(SICI)1098-1098(1997)8:6<519::AID-IMA4>3.0.CO;2-D' (cit. on p. 6.18).
noll:05:cpm	[115]	D. C. Noll, J. A. Fessler, and B. P. Sutton. "Conjugate phase MRI reconstruction with spatially variant sample density correction." In: <i>IEEE Trans. Med. Imag.</i> 24.3 (Mar. 2005), 325–36. DOI: 10.1109/TMI.2004.842452 (cit. on pp. 6.18, 6.27).
rasche:99:rod	[116]	V. Rasche et al. "Resampling of data between arbitrary grids using convolution interpolation." In: <i>IEEE Trans. Med. Imag.</i> 18.5 (May 1999), 385–92. DOI: 10.1109/42.774166 (cit. on p. 6.18).
schomberg:95:tgm	[117]	H. Schomberg and J. Timmer. "The gridding method for image reconstruction by Fourier transformation." In: <i>IEEE Trans. Med. Imag.</i> 14.3 (Sept. 1995), 596–607. DOI: 10.1109/42.414625 (cit. on p. 6.18).
schomberg:U2:nod	[118]	H. Schomberg. "Notes on direct and gridding-based Fourier inversion methods." In: <i>Proc. IEEE Intl. Symp. Biomed. Imag.</i> 2002, 645–8. DOI: 10.1109/ISBI.2002.1029340 (cit. on p. 6.18).
Chompson: 74.141	[119]	A. R. Thompson and R. N. Bracewell. "Interpolation and Fourier transformation of fringe visibilities." In: <i>The Astronomical Journal</i> 79.1 (Jan. 1974), 11–24. URL: http://adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=1974AJ7911T (cit. on p. 6.19).
osullivan:85:afs	[120]	J. D. O'Sullivan. "A fast sinc function gridding algorithm for Fourier inversion in computer tomography." In: <i>IEEE Trans. Med. Imag.</i> 4.4 (Dec. 1985), 200–7. DOI: 10.1109/TMI.1985.4307723 (cit. on p. 6.19).
slepian:64:psw	[121]	D. Slepian. "Prolate spheroidal wave functions, Fourier analysis and uncertainty, IV: Extensions to many dimensions; generalized prolate spheroidal functions." In: <i>Bell Syst. Tech. J.</i> 43.6 (Nov. 1964), 3009–57. DOI: 10.1002/j.1538-7305.1964.tb01037.x (cit. on p. 6.19).
slepian:83:sco	[122]	D. Slepian. "Some comments on Fourier analysis, uncertainty and modeling." In: <i>SIAM Review</i> 25.3 (July 1983), 379–93. DOI: 10.1137/1025078 (cit. on p. 6.19).
choi • 98 • aad	[123]	D. Rosenfeld. "An optimal and efficient new gridding algorithm using singular value decomposition." In: <i>Mag. Res. Med.</i> 40.1 (July 1998), 14–23. DOI: 10.1002/mrm.1910400103 (cit. on pp. 6.20, 6.28, 6.29).
sedarat:00:oto	[124]	H. Choi and D. C. Munson. "Analysis and design of minimax-optimal interpolators." In: <i>IEEE Trans. Sig. Proc.</i> 46.6 (June 1998), 1571–9. DOI: 10.1109/78.678470 (cit. on pp. 6.21, 6.27, 6.28, 6.29, 6.34).
	[125]	H. Sedarat and D. G. Nishimura. "On the optimality of the gridding reconstruction algorithm." In: <i>IEEE Trans. Med. Imag.</i> 19.4 (Apr. 2000), 306–17. DOI: 10.1109/42.848182 (cit. on pp. 6.21, 6.22, 6.23, 6.27).
calvagno:96:aid	[126]	G. Calvagno and D. C. Munson. "A frequency-domain approach to interpolation from a nonuniform grid." In: <i>Signal Processing</i> 52.1 (July 1996), 1–21. DOI: 10.1016/0165-1684 (96) 00058-8 (cit. on pp. 6.23, 6.28, 6.29).
snyder:92:dst	[127]	D. L. Snyder, T. J. Schulz, and J. A. O'Sullivan. "Deblurring subject to nonnegativity constraints." In: <i>IEEE Trans. Sig. Proc.</i> 40.5 (May 1992), 1143–50. DOI: 10.1109/78.134477 (cit. on p. 6.25).
vardi:93:fid	[128]	Y. Vardi and D. Lee. "From image deblurring to optimal investments: maximum likelihood solutions for positive linear inverse problems." In: <i>J. Royal Stat. Soc. Ser. B</i> 55.3 (1993), 569–612. URL: http://www.jstor.org/stable/2345873 (cit. on p. 6.25).
vard1:85:asm	[129]	Y. Vardi, L. A. Shepp, and L. Kaufman. "A statistical model for positron emission tomography." In: <i>J. Am. Stat. Assoc.</i> 80.389 (Mar. 1985), 8–37. URL: http://www.jstor.org/stable/2288030 (cit. on p. 6.25).
qian:02:rom	[130]	Y. Qian, J. Lin, and D. Jin. "Reconstruction of MR images from data acquired on an arbitrary k-space trajectory using the same-image weight." In: <i>Mag. Res. Med.</i> 48.2 (Aug. 2002), 306–11. DOI: 10.1002/mrm.10218 (cit. on p. 6.25).
dwork:17:als	[131]	N. Dwork et al. "A least squares optimal density compensation function for gridding." In: <i>Proc. Intl. Soc.</i> <i>Mag. Res. Med.</i> 2017, p. 4009. URL: http://archive.ismrm.org/2017/4009.html (cit. on p. 6.25).

ferrara:13:row

- [132] A. A. Samsonov, E. G. Kholmovski, and C. R. Johnson. "Determination of the sampling density compensation function using a point spread function modeling approach and gridding approximation." In: *Proc. Intl. Soc. Mag. Res. Med.* 2003, p. 477. URL: http://dev.ismrm.org//2003/0477.html (cit. on p. 6.25).
- [133] H. Tan and Y. Zheng. "Point spread function optimization for MRI reconstruction." In: *Proc. IEEE Conf. Acoust. Speech Sig. Proc.* Vol. 2. 2005, 477–80. DOI: 10.1109/ICASSP.2005.1415445 (cit. on p. 6.25).
  - [134] S. Inati et al. "Fast, optimal weighting for image reconstruction from arbitrary k-space trajectories." In: *Proc. Intl. Soc. Mag. Res. Med.* 2005, p. 2297. URL:

http://cds.ismrm.org/ismrm-2005/Files/02297.pdf (cit.on p. 6.27).

- [135] L. Lamalle. "Gridding: finite kernel support extent implies that sampling density compensation should be estimated locally." In: *ISMRM Workshop on Data Sampling and Image Reconstruction*. 2016 (cit. on p. 6.27).
  - [136] L. Lamalle. "Gridding: A highly parallel computing implementation using locally determined weights." In: *ISMRM Workshop on Data Sampling and Image Reconstruction*. 2016 (cit. on p. 6.27).
    - [137] J. G. Pipe. "Reconstructing MR images from undersampled data: Data-weighting considerations." In: *Mag. Res. Med.* 43.6 (June 2000), 867–75. DOI: 10.1002/1522-2594 (200006) 43:6<867::AID-MRM13>3.0.CO; 2-2 (cit. on p. 6.27).
    - [138] M. Ferrara, J. T. Parker, and M. Cheney. "Resolution optimization with irregularly sampled Fourier data." In:
    - *Inverse Prob.* 29.5 (May 2013), p. 054007. DOI: 10.1088/0266-5611/29/5/054007 (cit. on p. 6.27).
      - [139] D. Mitsouras et al. "Basis function cross-correlations for robust k-space sample density compensation, with application to the design of radiofrequency excitations." In: *Mag. Res. Med.* 57.2 (Feb. 2007), 338–52. DOI: 10.1002/mrm.21125 (cit. on p. 6.27).
    - [140] K. O. Johnson and J. G. Pipe. "Convolution kernel design and efficient algorithm for sampling density correction." In: *Mag. Res. Med.* 61.2 (Feb. 2009), 439–47. DOI: 10.1002/mrm.21840 (cit. on pp. 6.27, 6.32).
      - [141] J. L. Yen. "On nonuniform sampling of bandlimited signals." In: *IRE Trans. Circuit Theory* 3.4 (Dec. 1956), 251–7. DOI: 10.1109/TCT.1956.1086325 (cit. on p. 6.28).
- [142] F. J. Beutler. "Error free recovery of signals from irregularly spaced samples." In: *SIAM Review* 8.3 (July 1966), 328–35. DOI: 10.1137/1008065 (cit. on p. 6.28).
- [143] A. J. Jerri. "The Shannon sampling theorems, its various extensions and applications, a tutorial review." In: *Proc. IEEE* 65.11 (Nov. 1977), 1565–96. DOI: 10.1109/PROC.1977.10771 (cit. on p. 6.28).
  - [144] D. S. Chen and J. P. Allebach. "Analysis of error in reconstruction of two-dimensional signals from irregularly spaced samples." In: *IEEE Trans. Acoust. Sp. Sig. Proc.* 35.2 (Feb. 1987), 173–80. DOI: 10.1109/TASSP.1987.1165115 (cit. on pp. 6.28, 6.29).
  - [145] K. D. Sauer and J. P. Allebach. "Iterative reconstruction of band-limited images from nonuniformly spaced samples." In: *IEEE Trans. Circ. Sys.* 34.12 (Dec. 1987), 1497–506 (cit. on pp. 6.28, 6.30).
    - [146] P. L. Butzer and R. L. Stens. "Sampling theory for not necessarily band-limited function: A historical overview." In: *SIAM Review* 34.1 (Mar. 1992), 40–53. DOI: 10.1137/1034002 (cit. on p. 6.28).
    - [147] H. G. Feichtinger and K. Gröchenig. "Iterative reconstruction of multivariate band-limited functions from irregular sampling values." In: *SIAM J. Math. Anal.* 23.1 (Jan. 1992), 244–61. DOI: 10.1137/0523013 (cit. on p. 6.28).
    - [148] N. P. Willis and Y. Bresler. "Norm invariance of minimax interpolation." In: *IEEE Trans. Info. Theory* 38.3 (May 1992), 1177–81. DOI: 10.1109/18.135664 (cit. on pp. 6.28, 6.29).
- [149] D. J. Wingham. "The reconstruction of a band-limited function and its Fourier transform from a finite number of samples at arbitrary locations by singular value decomposition." In: *IEEE Trans. Sig. Proc.* 40.3 (Mar. 1992), 559–70. DOI: 10.1109/78.120799 (cit. on pp. 6.28, 6.29).
  - [150] A. Zakhor and G. Alvstad. "Two-dimensional polynomial interpolation from nonuniform samples." In: *IEEE Trans. Sig. Proc.* 40.1 (Jan. 1992), 169–80. DOI: 10.1109/78.157192 (cit. on p. 6.28).
  - [151] A. Aldroubi and H. Feichtinger. "Exact iterative reconstruction algorithm for multivariate irregularly sampled functions in spline-like spaces: the L<sub>p</sub>-theory." In: Proc. Amer. Math. Soc. 126.9 (Sept. 1998), 2677–86. DOI: 10.1090/S0002-9939-98-04319-6 (cit. on pp. 6.28, 6.30).

parks:78:imm

- [153] P. Marziliano and M. Vetterli. "Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations." In: *IEEE Trans. Sig. Proc.* 48.12 (Dec. 2000), 3462–71. DOI: 10.1109/78.887038 (cit. on p. 6.28).
- [154] A. Aldroubi and K. Gröchenig. "Nonuniform sampling and reconstruction in shift-invariant spaces." In: *SIAM Review* 43.4 (Dec. 2001), 585–620. DOI: 10.1137/S0036144501386986 (cit. on p. 6.28).
  - [155] F. Marvasti. *Nonuniform sampling: Theory and practice*. New York: Kluwer, 2001. URL: http://www.wkap.nl/prod/b/0-306-46445-4 (cit. on p. 6.28).
    - [156] E. Meijering. "A chronology of interpolation: from ancient astronomy to modern signal and image processing." In: *Proc. IEEE* 90.3 (Mar. 2002), 319–42. DOI: 10.1109/5.993400 (cit. on p. 6.28).
  - T. W. Parks and D. P. Kolba. "Interpolation minimizing maximum normalized error for band-limited signals." In: *IEEE Trans. Acoust. Sp. Sig. Proc.* 26.4 (Aug. 1978), 381–4. DOI: 10.1109/TASSP.1978.1163103 (cit. on pp. 6.28, 6.29).
    - [158] D. A. Fish, J. Grochmalicki, and E. R. Pike. "Scanning SVD method for restoration of images with space-variant blur." In: J. Opt. Soc. Am. A 13.3 (Mar. 1996), 464–9. DOI: 10.1364/JOSAA.13.000464 (cit. on p. 6.29).
- [159] D. Rosenfeld. "New approach to gridding using regularization and estimation theory." In: *Mag. Res. Med.* 48.1 (July 2002), 193–202. DOI: 10.1002/mrm.10132 (cit. on p. 6.29).
  - [160] X. Jiang and D. Thedens. "New iterative gridding algorithm using conjugate gradient method." In: *Proc. SPIE 5370 Medical Imaging: Image Proc.* 2004, 1852–61. DOI: 10.1117/12.535685 (cit. on p. 6.29).
  - [161] A. S. Fahmy, B. Tawfik, and Y. M. Kadah. "Image reconstruction using shift-variant resampling kernel for magnetic resonance imaging." In: *Proc. SPIE 4684 Medical Imaging: Image Proc.* 2002, 825–33. DOI: 10.1117/12.467230 (cit. on p. 6.29).
- [162] S. Kashyap and M. Jacob. "A fast & accurate non iterative algorithm for regularized non-Cartesian MRI." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2010, 265–8. DOI: 10.1109/ISBI.2010.5490362 (cit. on p. 6.29).
- [163] S. Kashyap, Z. Yang, and M. Jacob. "Non-iterative regularized reconstruction algorithm for non-Cartesian MRI: NIRVANA." In: *Mag. Res. Im.* 29.2 (Feb. 2011), 222–9. DOI: 10.1016/j.mri.2010.08.017 (cit. on p. 6.29).
  - [164] F. J. Beutler. "Sampling theorems and bases in a Hilbert space." In: Inform. and Control 4.2-3 (Sept. 1961), 97–117. URL: http://theory.lcs.mit.edu/~iandc/ic61.html (cit. on p. 6.29).
- [165] L. Levi. "Fitting a bandlimited signal to given points." In: *IEEE Trans. Info. Theory* 11.3 (July 1965), 372–6.
   DOI: 10.1109/TIT.1965.1053777 (cit. on p. 6.29).
  - [166] H. Landau. "Extrapolating a band-limited function from its samples taken in a finite interval." In: *IEEE Trans. Info. Theory* 32.4 (July 1986), 464–70. DOI: 10.1109/TIT.1986.1057205 (cit. on p. 6.29).
    - [167] S. Yeh and H. Stark. "Iterative and one-step reconstruction from nonuniform samples by convex projections." In: J. Opt. Soc. Am. A 7.3 (Mar. 1990), 491–9. DOI: 10.1364/JOSAA.7.000491 (cit. on pp. 6.29, 6.30).
- [168] J. L. Brown. "Anharmonic approximation and bandlimited signals." In: Inform. and Control 10.4 (Apr. 1967), 409-18. URL: http://theory.lcs.mit.edu/~iandc/ic67.html#Brown1967:409 (cit. on p. 6.29).
  - [169] N. T. Gaarder. "A note on multidimensional sampling theorem." In: *Proc. IEEE* 60.2 (Feb. 1972), 247–8.
     DOI: 10.1109/PROC.1972.8615 (cit. on p. 6.29).
    - [170] E. J. Diethorn and D. C. Munson. "A linear, time-varying system framework for noniterative discrete-time band-limited signal extrapolation." In: *IEEE Trans. Sig. Proc.* 39.1 (Jan. 1991), 55–68. DOI: 10.1109/78.80765 (cit. on p. 6.29).
    - [171] K. Miller. "Least-squares methods for ill-posed problems with a prescribed bound." In: SIAM J. Math. Anal.
       1.1 (Feb. 1970), 52–70. DOI: 10.1137/0501006 (cit. on p. 6.29).

kolba:83:oef

KOIDA.03.0EI	[172]	D. Kolba and T. Parks. "Optimal estimation for band-limited signals including time domain considerations." In: <i>IEEE Trans. Acoust. Sp. Sig. Proc.</i> 31.1 (Feb. 1983), 113–22. DOI: 10.1109/TASSP.1983.1164011 (cit. on p. 6.29).
wiley:78:rob	[173]	R. G. Wiley. "Recovery of bandlimited signals from unequally spaced samples." In: <i>IEEE Trans. Comm.</i> 26.1 (Jan. 1978), 135–7. DOI: 10.1109/TCOM.1978.1093962 (cit. on p. 6.30).
andberg: 94.tro	[174]	F. Marvasti, M. Analoui, and M. Gamshadzahi. "Recovery of signals from nonuniform samples using iterative methods." In: <i>IEEE Trans. Sig. Proc.</i> 39.4 (Apr. 1991), 872–8. DOI: 10.1109/78.80909 (cit. on p. 6.30).
	[175]	I. W. Sandberg. "The reconstruction of band-limited signals from nonuniformly spaced samples." In: <i>IEEE Trans. Circ. Sys. I, Fundamental theory and applications</i> 41.1 (Jan. 1994), 64–6. DOI: 10.1109/81.260224 (cit. on p. 6.30).
ntinger:93:Iir	[176]	H. G. Feichtinger and T. Strohmer. "Fast iterative reconstruction of band-limited images from non-uniform sampling values." In: <i>CAIP'93</i> . 1993, 82–9. URL:
gabr:06:dig	[177]	R. E. Gabr et al. "Deconvolution-interpolation gridding (DING): Accurate reconstruction for arbitrary k-space trajectories." In: <i>Mag. Res. Med.</i> 56.6 (Dec. 2006), 1182–91. DOI: 10.1002/mrm.21095 (cit. on p. 6.30).
iperwas:16:tsa	[178]	A. Kiperwas, D. Rosenfeld, and Y. C. Eldar. <i>The SPURS algorithm for resampling an irregularly sampled signal onto a cartesian grid</i> . 2016. URL: http://arxiv.org/abs/1603.04726 (cit. on p. 6.30).
beatty:05:rgr	[179]	P. J. Beatty, D. G. Nishimura, and J. M. Pauly. "Rapid gridding reconstruction with a minimal oversampling ratio." In: <i>IEEE Trans. Med. Imag.</i> 24.6 (June 2005), 799–808. DOI: 10.1109/TMI.2005.848376 (cit. on pp. 6.31, 6.35, 6.37, 6.41).
esterle:99:srb	[180]	C. Oesterle et al. "Spiral reconstruction by regridding to a large rectilinear matrix: A practical solution for routine systems." In: <i>J. Mag. Res. Im.</i> 10.1 (July 1999), 84–92. DOI: 10.1002/(SICI)1522-2586(199907)10:1<84::AID-JMRI12>3.0.CO; 2-D (cit. on p. 6.31).
riguchi:04:inn	[181]	H. Moriguchi and J. L. Duerk. "Iterative next-neighbor regridding (INNG): Improved reconstruction from nonuniformly sampled k-space data using rescaled matrices." In: <i>Mag. Res. Med.</i> 51.2 (Feb. 2004), 343–52. DOI: 10.1002/mrm.10692 (cit. on p. 6.31).
gai:13:mia	[182]	J. Gai et al. "More IMPATIENT: A gridding-accelerated Toeplitz-based strategy for non-Cartesian high-resolution 3D MRI on GPUs." In: <i>J. of Parallel and Distributed Computing</i> 73.5 (May 2013), 686–97. DOI: 10.1016/j.jpdc.2013.01.001 (cit. on p. 6.32).
pipe:95:ara	[183]	J. G. Pipe and J. L. Duerk. "Analytical resolution and noise characteristics of linearly reconstructed magnetic resonance data with arbitrary k-space sampling." In: <i>Mag. Res. Med.</i> 34.2 (Aug. 1995), 170–8. DOI: 10.1002/mrm.1910340207 (cit. on p. 6.34).
adcock:14:osr	[184]	B. Adcock, M. Gataric, and A. Hansen. "On stable reconstructions from nonuniform Fourier measurements." In: <i>SIAM J. Imaging Sci.</i> 7.3 (2014), 1690–1723. DOI: 10.1137/130943431 (cit. on p. 6.34).
slepian:78:psw	[185]	D. Slepian. "Prolate spheroidal wave functions. Fourier analysis and uncertainty, V: The discrete case." In: <i>Bell Syst. Tech. J.</i> 57.5 (May 1978), 1371–430. DOI: 10.1002/j.1538-7305.1978.tb02104.x (cit. on p. 6.36).
mathews:85:tdp	[186]	J. D. Mathews, J. K. Breakall, and G. K. Karawas. "The discrete prolate spheroidal filter as a digital signal-processing tool." In: <i>IEEE Trans. Acoust. Sp. Sig. Proc.</i> 33.6 (Dec. 1985), 1471–8. DOI: 10.1109/TASSP.1985.1164732 (cit. on p. 6.36).
fessler:07:onb	[187]	J. A. Fessler. "On NUFFT-based gridding for non-Cartesian MRI." In: J. Mag. Res. 188.2 (Oct. 2007), 191–5. DOI: 10.1016/j.jmr.2007.06.012 (cit. on p. 6.37).
lewitt:90:mdi	[188]	R. M. Lewitt. "Multidimensional digital image representations using generalized Kaiser-Bessel window functions." In: <i>J. Opt. Soc. Am. A</i> 7.10 (Oct. 1990), 1834–46. DOI: 10.1364/JOSAA.7.001834 (cit. on p. 6.37).
lewitt:92:atv	[189]	R. M. Lewitt. "Alternatives to voxels for image representation in iterative reconstruction algorithms." In: <i>Phys. Med. Biol.</i> 37.3 (Mar. 1992), 705–16. DOI: 10.1088/0031-9155/37/3/015 (cit. on p. 6.37).
wajer:99:sef	[190]	F. T. A. W. Wajer et al. "Simple equation for optimal gridding parameters." In: <i>Proc. Intl. Soc. Mag. Res.</i> <i>Med.</i> 1999, p. 663. URL: http://cds.ismrm.org/ismrm-1999/PDF3/663.PDF (cit. on p. 6.37).

- [191] M. Jacob. "Optimized least-square nonuniform fast Fourier transform." In: *IEEE Trans. Sig. Proc.* 57.6 (June 2009), 2165–77. DOI: 10.1109/TSP.2009.2014809 (cit. on p. 6.37).
- [192] Z. Yang and M. Jacob. "Mean square optimal NUFFT approximation for efficient non-Cartesian MRI reconstruction." In: J. Mag. Res. 242 (May 2014), 126–35. DOI: 10.1016/j.jmr.2014.01.016 (cit. on p. 6.37).
  - [193] L. Sha, H. Guo, and A. W. Song. "An improved gridding method for spiral MRI using nonuniform fast Fourier transform." In: *J. Mag. Res.* 162.2 (June 2003), 250–8. DOI: 10.1016/S1090-7807 (03) 00107-1 (cit. on pp. 6.39, 6.40).
  - [194] J. A. Fessler et al. "Toeplitz-based iterative image reconstruction for MRI with correction for magnetic field inhomogeneity." In: *IEEE Trans. Sig. Proc.* 53.9 (Sept. 2005), 3393–402. DOI: 10.1109/TSP.2005.853152 (cit. on p. 6.43).
- [195] G. McGibney et al. "Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI." In: *Mag. Res. Med.* 30.1 (July 1993), 51–9. DOI: 10.1002/mrm.1910300109 (cit. on p. 6.43).
- [196] J. A. Fessler and D. C. Noll. "Iterative image reconstruction in MRI with separate magnitude and phase regularization." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2004, 209–12. DOI: 10.1109/ISBI.2004.1398511 (cit. on pp. 6.43, 6.44).
  - [197] M. Bydder and M. D. Robson. "Partial Fourier partially parallel imaging." In: *Mag. Res. Med.* 53.6 (June 2005), 1393–401. DOI: 10.1002/mrm.20492 (cit. on p. 6.44).
- [198] M. L. Lauzon and B. K. Rutt. "Effects of polar sampling in k-space." In: *Mag. Res. Med.* 36.6 (Dec. 1996), 940–9. DOI: 10.1002/mrm.1910360617 (cit. on pp. 6.44, 6.45).
  - [199] P. M. Joseph. "Sampling errors in projection reconstruction MRI." In: *Mag. Res. Med.* 40.3 (Sept. 1998), 460–6. DOI: 10.1002/mrm.1910400317 (cit. on p. 6.44).
- [200] M. Fenn, S. Kunis, and D. Potts. "On the computation of the polar FFT." In: *Applied and Computational Harmonic Analysis* 22.2 (Mar. 2007), 257–63. DOI: 10.1016/j.acha.2006.05.009 (cit. on p. 6.44).
- [201] D. Potts and G. Steidl. "Fast summation at nonequispaced knots by NFFT." In: *SIAM J. Sci. Comp.* 24.6 (2003), 2013–37. DOI: 10.1137/S1064827502400984 (cit. on p. 6.45).