

Signal Processing Algorithms for Simplified-Hardware
Multistatic Synthetic Aperture Radar

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Jonathan Darrel Coker

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

Ahmed H. Tewfik, Advisor

October 2008

Acknowledgments

When a student reaches a certain age, the onus of completing an advanced graduate program falls not only upon the student, but also on all those to whom he is obligated in duty and in affection. I particularly wish to thank Dr. Barry Gilbert of the Mayo Clinic for his encouragement at the beginning of this effort; my friend and colleague Michael J. Ross at Hitachi GST was steadfast in his support at its end. Dr. Ahmed Tewfik's skilful teaching and singularly profound advice will long be an inspiration to me; his is a valued friendship which I waste no space by acknowledging.

Each of these benefactors would freely admit, however, that the greatest debt is payable to my family Kay, Drew, Bren, and Dara; who so generously picked up the slack and kept the family life vibrant despite my divided attentions during this time.

Abstract

This work represents contributions to the state of the art of radar tomography and SAR in three main areas. First, we develop a SAR measurement and image reconstruction framework for a downlooking, omnidirectional and multistatic SAR system. This framework, a significant generalization of previous work for downlooking monostatic SARs, models the SAR measurement process as a generalized integral transform of a reflectivity function.

Secondly, based on this measurement and inversion framework, we resolve the novel question of SAR vehicle trajectory design for the purpose of optimal image quality in the presence of measurement noise. In addition, we introduce the concept of a *random location* multistatic trajectory, and show that (for large enough problems) such a strategy is likely to approximate an upper performance bound; therefore a random strategy is likely to perform as well or better than any deterministic trajectory strategy.

Finally, we expose connections between the Welton phase-coded radar sequences and the Discrete Wavelet transform. This connection directly yields very simple and natively high-speed implementations of pulse compression filters at the radar receiver. Furthermore, we generalize the idea of zero correlation zone sequences by developing methods to generate transmit waveforms which extend the perfect correlation width of such filter outputs without modifying the simplicity of the receive filters. This method provides a new system tradeoff between transmit energy and receiver hardware complexity and power.

Contents

1	Foundations	1
1.1	Main contributions	1
1.2	Introduction	2
1.3	Relevant topics in integral geometry	4
1.4	Medical tomography	7
1.5	Traditional SAR technology	9
1.6	Radar tomography	15
1.7	Thesis summary	18
2	An elliptical formulation for multistatic omnidirectional SAR	20
2.1	A proposed bistatic SAR geometry	21
2.1.1	An elliptical Radon transform	22
2.1.2	General integral transforms	28
2.2	Backprojection operators for \mathcal{R}_e	30
2.2.1	Frequency-domain backprojection	32
2.2.2	Discussion	35
2.3	The adjoint of \mathcal{R}_e	39
2.4	Numerical backprojection results	43
2.5	The approximate inverse method	45
2.5.1	Maximum likelihood image reconstruction	48
2.6	Numerical approximate inverse results	53
2.7	Conclusions	56
3	Performance synthesis for UAV trajectories in multistatic SAR	58
3.1	A noise figure performance metric	61
3.2	Absolute performance bounds	66
3.2.1	Conditions for meeting the lower bound	67
3.3	Pragmatic performance bounds	68
3.3.1	Estimates of the dominant eigenvalue	69
3.3.2	An improved lower bound for finite σ_λ^2	69
3.4	Physical interpretation of performance bounds	70
3.4.1	Preliminaries	71
3.4.2	Bounds in terms of trajectory parameters	74

3.5	Guidelines for trajectory design	75
3.6	Application to example trajectories	76
3.6.1	A standard bistatic trajectory	76
3.6.2	A degenerate case	77
3.6.3	A merry-go-round multistatic trajectory	77
3.6.4	Random trajectories	78
3.6.5	Performance survey results	78
3.6.6	Detailed comparison of random and standard trajectories . . .	81
3.7	Performance estimation by random matrix theory	85
3.7.1	System scaling equations	87
3.8	Edge effects and variable resolution imaging	88
3.9	Random-location sampling and compressed sensing	92
3.10	Conclusions	94
4	Simplified-hardware ranging systems using wavelet decomposition	96
4.1	Background	99
4.1.1	Sequence constructions	102
4.1.2	DWT implementations	104
4.2	Generalized zero autocorrelation zone extensions	106
4.2.1	Generalized extensions	107
4.2.2	Implementation	113
4.3	Efficient use of the second output	114
4.4	Conclusions	117
5	Final remarks and future work	118
	Bibliography	120

List of Tables

3.1	Numerical comparison of SAR performance and bounds	80
-----	--	----

List of Figures

1.1	Parallel-beam tomography: line integrals of a function $f(x, y)$ versus of the offset r at a particular direction $\theta = -\frac{\pi}{4}$ (left). A density plot of the transform $\mathcal{R}[f]$ as a function of the offset r and direction θ (right).	8
1.2	Traditional SAR geometry for a UAV moving above the x axis and a point reflector in the (x, y) plane at a slant range r_1 from the flight trajectory.	10
1.3	Normalized range/azimuth ambiguity function for an X-band (10 GHz) SAR system of velocity 250 meters/second with a point reflector at range $r_1 = 1$ km, azimuth $x_1 = 0$ meters, Doppler bandwidth $= \pm 10$ kHz.	14
1.4	Illustration of the connection between spotlight-mode SAR processing and computed tomography. (Figure inspired by J. A. Richards.)	16
2.1	Geometrical specification for a single transmit/receive bistatic pair over a flat plane.	23
2.2	The point reflector response $\mathcal{R}_e[\delta(x_0, y_0)]$ for $x_0 = 0$, $y_0 = 2$, $h = 3$, and $d = 2$.	27
2.3	Geometrical depiction of a simple multistatic configuration over a flat plane.	29
2.4	The transfer functions W_0 and W_2 versus R using $d = 1$, $h = 2$, and various values of ϕ .	37
2.5	The transfer function W_A versus r using $d = 2$, $h = 3$, and various values of $\cos \theta = x/r$.	42
2.6	Unfiltered backprojections using various weighting strategies for a Shepp-Logan phantom, derived over an aperture of ± 2.5 image widths.	44
2.7	Comparison of 75×75 approximate inverse reconstructions of the \mathcal{R}_e trajectory without (left) and with (right) function-approximation error.	54
2.8	Comparison of 75×75 approximate inverse reconstructions of the \mathcal{R}_e trajectory (left) and a random multistatic trajectory (right). Both measurements exhibit finite function-approximation error.	56

3.1	A representative example of four ellipses of integration, each representing a row of the matrix \mathbf{A} . The light dotted curves represent contours of some reflectivity function $f(x, y)$	62
3.2	Nominal relationships between the sum of squares of a row of \mathbf{A} and the associated ellipse length, for different values of n	72
3.3	A merry-go-round multistatic trajectory with 9 SAR vehicles over a square reflectivity plane.	79
3.4	The empirical singular value distributions of \mathbf{A} for $\mathcal{R}_e f$ and for a random sampling trial, with $m \approx 12000$	82
3.5	Noise gain of deterministic and random sampling as a function of the measurement size m and constant $n = 50 \times 50$	83
3.6	Comparison of reconstructed images with the same measurement noise power, using $m \approx 9000$	84
3.7	Comparison of G estimates and actual performance for an $n = 75 \times 75$ image using random trajectories and various values of m	86
3.8	The effect on pixel energy of deleting out-of-bounds sampling ellipses in an $n = 50 \times 50$ merry-go-round trajectory. (An image width = 2 distance units).	89
3.9	Superpixel pattern of dimensions 8, 4, 2, and 1, from light to dark (left). Noiseless reconstruction with no reflectors in superpixel locations (middle). Noiseless reconstruction with bright single-pixel reflector within the marked superpixel location (right).	91
3.10	Noiseless minimum- ℓ_2 norm reconstruction (left) and minimum-TV reconstruction (right) for a $75 \times 75 = 5625$ pixel image and $m = 1765$ random-location SAR measurements.	94
4.1	A forward discrete wavelet transform.	100
4.2	A polyphase implementation of a discrete wavelet transform.	101
4.3	DWT decomposition hardware for \mathbf{H} with $L = 32$	105
4.4	Extending the sequence \mathbf{c} for perfect correlation filter outputs.	107
4.5	The transmit energy costs versus designed zero correlation zone τ for all sequences in \mathbf{W}_7	111
4.6	Generalized extensions for the first sequence of \mathbf{W}_7 up to $\tau = 63$	112
4.7	An efficient pipelined half-rate implementation of the matched filter of a sequence \mathbf{c}	113
4.8	DWT decomposition of \mathbf{e} with $L = 2^{12}$ for different point reflector locations.	115
4.9	Normalized maximum sidelobes for DWT decomposition of \mathbf{e} versus sequence length.	116

Chapter 1

Foundations

1.1 Main contributions

This work represents contributions to the state of the art of radar tomography and SAR in three main areas. First, we develop a SAR measurement and image reconstruction framework for a downlooking, omnidirectional and multistatic SAR system. This framework, a significant generalization of previous work for downlooking monostatic SARs, models the SAR measurement process as a generalized integral transform of a reflectivity function.

Secondly, based on this measurement and inversion framework, we resolve the novel question of SAR vehicle trajectory design for the purpose of optimal image quality in the presence of measurement noise. In addition, we introduce the concept of a *random location* multistatic trajectory, and show that (for large enough problems) such a strategy is likely to approximate an upper performance bound; therefore a random strategy is likely to perform as well or better than any deterministic trajectory strategy.

Finally, we expose connections between the Welfi phase-coded radar sequences

(known since the 1960's) and the Discrete Wavelet transform. This connection directly yields very simple and natively high-speed implementations of pulse compression filters at the radar receiver. Furthermore, we generalize the idea of zero correlation zone sequences by developing methods to generate transmit waveforms which extend the perfect correlation width of such filter outputs without modifying the simplicity of the receive filters. This method provides a new system tradeoff between transmit energy and receiver hardware complexity and power.

1.2 Introduction

Synthetic aperture radar (SAR) is a means by which a scene image may be reconstructed by examining photonic test signals which have been reflected from objects in the scene. One simple reconstruction method, from a signal processing point of view, is to illuminate one small piece of the scene at a time using narrow-beamwidth transmissions. The brightness or darkness of a corresponding image pixel is then determined by a measurement of the strength of any reflections appearing at some conveniently-placed receive antenna.

However, from a system perspective, such a reconstruction method is impractical because antennas of physically large size (“aperture”) and high complexity are required to achieve a tightly focused beam [1]. In synthetic aperture imaging, the resolution advantages of a large antenna aperture are effectively *synthesized* using relatively small antennas which are moved about relative to the scene in some way. In such a case, the measurements at a receive antenna are no longer due to reflections from the minimum area corresponding to the image resolution, but are measurements of average reflection strength taken over a larger area. This is SAR signal processing in a nutshell: to reconstruct information about a scene’s reflectivity from such

measurements [2].

SAR imaging technology is now exhibited in a rich variety of modalities, which depend on the system geometry, the trajectory of the antennas, and many other parameters. The fundamental classification of any SAR system is the configuration of its antennas. Most traditional SAR systems are *monostatic*, that is, the transmit and receive antennas are substantially collocated at all times [3].

In the present work, *bistatic* and *multistatic* configurations are examined, in which multiple transmit and receive antennas may exist at distinct locations. The defining strategy of this work is to define signal processing algorithms which support low-complexity, low-power hardware on the moving SAR platforms. If such platforms are inexpensive enough, then the system may gain significant advantages in image resolution, exposure time, logistics, or area coverage by using many platforms. We shall not necessarily require low computational complexity at the central processing location for the final formation of an image. A main goal of this research will be to examine the system tradeoffs of such a strategy.

The intended application of this work is in support of multiple Unmanned Airborne Vehicle (UAV) systems for SAR imaging. In such systems, the power consumption of the SAR electronics is of critical importance because it trades off directly with the available time aloft. On the other hand, such systems may further require that (at least) *some* initial processing be done on board, to enable some sort of data compression for efficient transmission to the processing unit. It is generally too costly in a high-resolution, high-bandwidth UAV system to simply retransmit the received waveform to a centralized location. In many situations, such as covert imaging, the amount of data sent from the UAVs to the processing unit should be minimized as much as possible. These conflicting requirements provide further focus on the problem of simplified local signal processing algorithms in this context.

The literature supporting SAR research is wide ranging, multidisciplinary, and continually expanding; furthermore, SAR itself is a small subset of a large number of optical imaging techniques; which in turn are part of the general study of *remote sensing*, which may be said to include sonar imaging, ground imaging, geophysical and astrophysical measurements, and any of a large number of tomographies used in medical applications.

The following sections provide a survey of the several mathematical and engineering disciplines which have provided the foundation for SAR processing as it exists today. These topics include integral geometry, traditional SAR processing, medical tomography, and the emerging discipline of radar tomography.

1.3 Relevant topics in integral geometry

In an oft-cited and, according to Madych [4], widely-unread 1917 monograph, Radon [5] introduced an integral transform in which a scalar function in \mathbb{R}^n is expressed by its integrals over $(n - 1)$ -dimensional hyperplanes. Radon developed inversion formulas for his transform, which almost never reoccur directly in the literature. Madych [4] notes this fact, and quotes one of Radon's results for a scalar function $f(x)$ with $x \in \mathbb{R}^2$, in modern notation, as follows.

If a transform \mathcal{R} of a function f is defined as

$$\mathcal{R}[f](\theta, t) = \int_{-\infty}^{\infty} f(tu_{\theta} + sv_{\theta}) ds \quad (1.1)$$

where $u_{\theta} = (\cos \theta, \sin \theta)$ and $v_{\theta} = (-\sin \theta, \cos \theta)$, then under several technical condi-

tions on f , $f(x)$ may be recovered from $\mathcal{R}[f](\theta, t)$ by the following procedure:

$$F_x(t) = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{R}[f](\theta, \langle x, u_\theta \rangle + t) d\theta, \quad (1.2)$$

$$f(x) = -\frac{1}{\pi} \int_0^\infty \frac{dF_x(t)}{t} \quad (1.3)$$

where $\langle x, u_\theta \rangle = x_1 \cos \theta + x_2 \sin \theta$ is the scalar product of x and u_θ .

Natterer [6] quotes Radon's result in a more commonly-used form, now generally thought of as the traditional \mathbb{R}^2 Radon transform and its inverse:

$$\mathcal{R}[f](\theta, s) = \int_{x \cdot \theta = s} f(x) dx = g(\theta, s) \quad (1.4)$$

$$f(x) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_{\mathbb{R}^1} \frac{g'(\theta, s)}{x \cdot \theta - s} d\varphi ds, \quad (1.5)$$

where g' is the derivative of g with respect to s , $s \in \mathbb{R}^1$, and $\theta = (\cos \varphi, \sin \varphi)^T$. This transform is also known as the *x-ray transform* in two dimensions. The forward transform \mathcal{R} generates the line integrals of $f(x)$ over all straight lines described by the direction θ and offset from the origin s . The result states, for example, that the shape of a two dimensional object can be uniquely reconstructed from the length of all line segments through the object. More generally, the result indicates that a two dimensional function can be recovered from its line integrals.

Radon's work itself did not arise *ex nihilo*. Another German, P. Funk, showed in 1913 that an even function defined on the unit sphere \mathbb{S}^2 can be reconstructed from its line integrals over all great circles on the sphere [7], thus predating Radon's paper by several years. This result is now called the Funk transform. Very recently, Yarman [8, 9] mapped a current SAR inversion problem into a form invertible via the Funk transform.

In 1955, Fritz John extended the ideas of Radon and Funk in a beautiful little

book [10] and applied them to solve various forms of partial differential equations. He anticipated many of the results reported 40 years later in other contexts. John appears to have coined the phrase *spherical means* to describe the *area* integrals of a function over spheres in \mathbb{R}^n , and he introduces the problem of recovering a function from its means over spheres of fixed radius r . It appears that in his application he considered surface integrals, not line integrals. Nevertheless, many researchers in tomography continue to refer to his work.

Previously, Funk had realized that the functions possessing odd symmetry between the upper and lower hemispheres are in the null space of his transform. Following Funk, John also pointed out that such problems do not always have unique inverse solutions. His simplest example is the spherical means of a function in \mathbb{R}^1 :

$$g(s) = \int_{-r}^r f(s+t)dt \quad (1.6)$$

for some fixed r . Any zero-mean function f of period $2r$ satisfies this equation with $g(s) = 0$. The identification of the null space of any of these transforms remains a common theme throughout their history.

The work of Cormack in the early 1960's [11, 12] will be mentioned later, in the context of the invention of computed tomography. In his successful quest to determine a function from its line integrals, he considered not only line integrals over straight lines, but also line integrals over circles. He may be the first to consider the use of circular contours to do so. Later, Cormack again contributed to generalized Radon transform theory by considering integrals of two-dimensional functions over two types of generalized curves [13, 14].

During the 1980's, mathematicians weighed in on the more formal definitions of generalized Radon transforms and on problems relative to the existence of solutions

and their uniqueness [15, 16, 17]. Fawcett and Andersson [18, 19] explored various extensions of the Radon transform in which the line integrals were taken over circles in the plane. Zalcman began several threads of Radon transform discussion in a recreational article entitled *Offbeat Integral Geometry* [20]. Discussion of the extensions of the classical Pompeiu problem by Laquer [21] explored the nature of the null space of particular Radon transforms. (The original Pompeiu problem asks: must a two-dimensional continuous function be zero if its integrals over all disks of a fixed radius are zero.)

The mathematics of generalized Radon transforms remains a lively research topic. In a recent and monumental work, Ehrenpreis [22] explores the connections between differential equations, generalized linear and nonlinear Radon transforms, differential geometry, integral geometry, and many results in modern analysis.

1.4 Medical tomography

In 1964, Cormack [11, 12] started a medical revolution by re-discovering the inverse of a Radon transform and applying it to reconstruct internal images of a material by non-invasive measurements taken from outside the body. He recognized that x-ray absorption measurements represent the line integrals of the absorption function of the material (see Figure 1.1). The figure shows how line integrals of a two dimensional object can be measured by shooting parallel x-ray beams through the object, and observing the intensity $\mathcal{R}[f](\theta = -\frac{\pi}{4}, r)$ at various offsets r by a string of x-ray detectors. The apparatus (or the object itself) can be rotated through all angles θ to sample the entire forward transform $\mathcal{R}[f](\theta, r)$. Furthermore, Cormack developed an approximate numerical method to reconstruct the two-dimensional absorption function from these measurements, taken at a finite number of angles. This brilliant

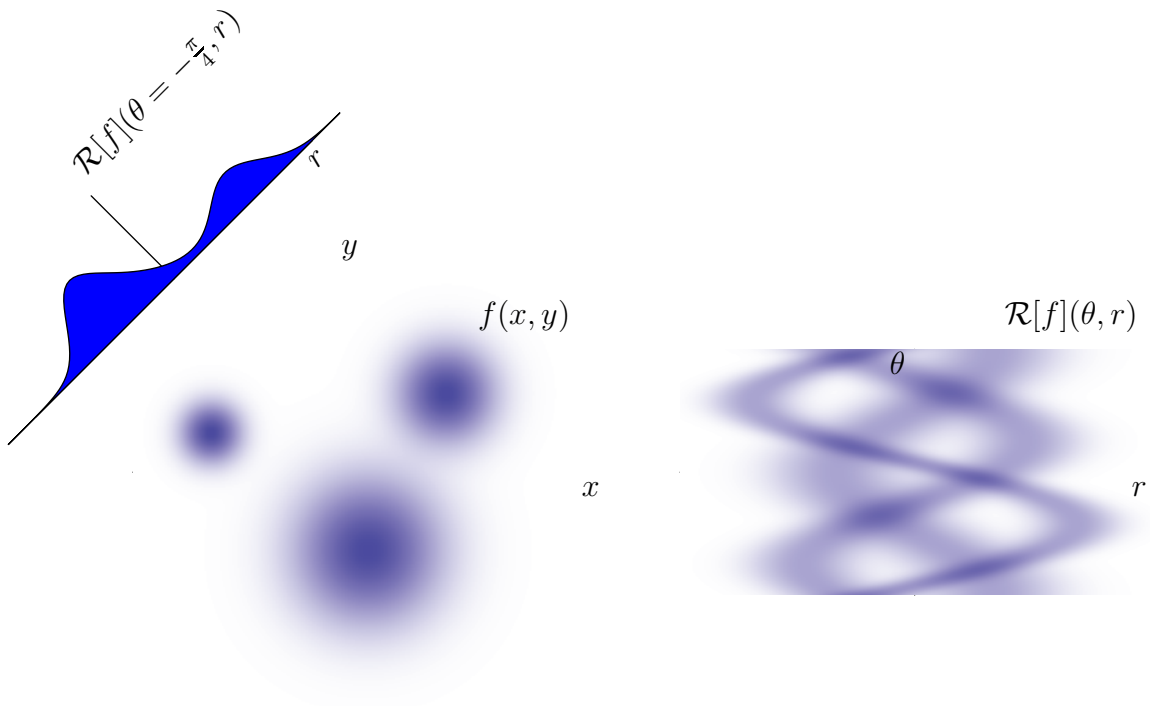


Figure 1.1: Parallel-beam tomography: line integrals of a function $f(x, y)$ versus of the offset r at a particular direction $\theta = -\frac{\pi}{4}$ (left). A density plot of the transform $\mathcal{R}[f]$ as a function of the offset r and direction θ (right).

combination of ideas revolutionized medical imaging practice. Cormack eventually shared a Nobel prize for the invention of computed tomography (CT).

Throughout the 1970's, efficient inversion schemes of the x-ray Radon transform were developed. These are nicely summarized in a survey article by Strichartz in 1982 [23, 24]. The well-known projection-slice theorem connected the Fourier transform with the Radon transform, and justified the commonly-used filtered-backprojection method, in which it was recognized that simply smearing the line integral uniformly back over its support almost gives the correct inversion for the x-ray transform. The fully-correct answer requires a two-dimensional deconvolution with a mollifier-like function called a Riesz kernel. Backprojection remains a key inversion technique in many types of inverse problems. The inversion consequences of more-easily-controlled x-ray sources (such as fan-beam sources) were studied and solved.

Other generalizations, including helical scans for three-dimensional tomography, efficient sampling techniques, and quasi-dynamic methods for handling patient movement, are among the active problems in medical tomography today.

1.5 Traditional SAR technology

The acronym *SAR*, since its inception in the 1960s, has now grown to encompass a broad array of signal processing techniques and hardware [3]. All versions of SAR generally use measurements of photonic signals reflected off a scene to generate images of the scene. It is therefore instructive to demonstrate the main concepts of a particularly elegant version of SAR: the original version.

The original SAR technologies may be described in the following way. As shown in Figure 1.2, a monostatic SAR vehicle travels above the x axis at a velocity v , continuously illuminating a point reflector in the (x, y) plane at $x = x_1$ with a

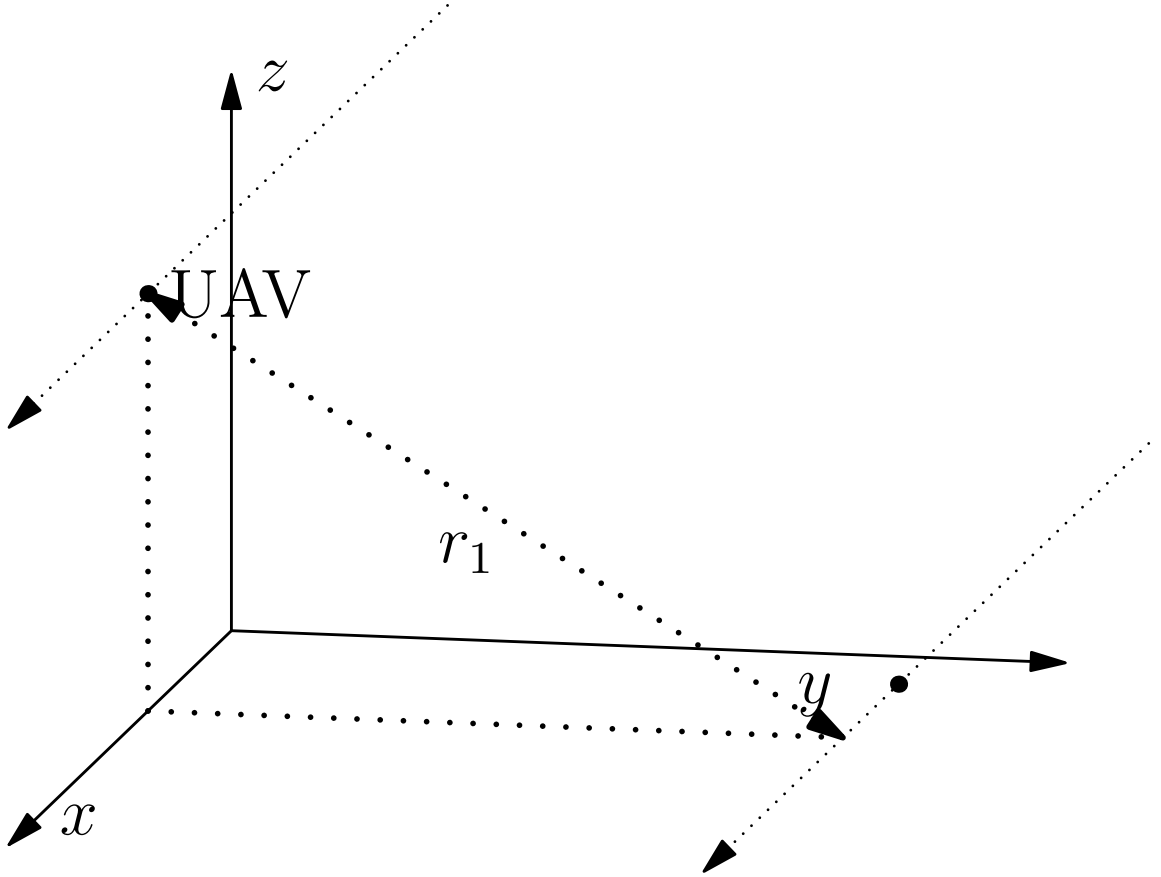


Figure 1.2: Traditional SAR geometry for a UAV moving above the x axis and a point reflector in the (x, y) plane at a slant range r_1 from the flight trajectory.

sinusoidal waveform of frequency ω . Following the derivations and approximations made in Bracewell [25], the return signal received at the SAR vehicle due to the point reflector is

$$s_1(t) = \sigma_1 \cos \left\{ \omega t - \frac{4\pi r_1}{\lambda} - \frac{2\pi}{\lambda r_1} (vt - x_1)^2 + \phi_1 \right\} \quad (1.7)$$

where σ_1 is an amplitude factor dependent on the path loss, transmitted power, and target reflectivity; $\lambda = \frac{2\pi c}{\omega}$ is the wavelength of the transmitted sinusoid; and ϕ_1 is some phase constant dependent on the complex reflectivity of the target. The

expression in equation 1.7 is valid only when the SAR is near the point reflector. In practice, the function $s_1(t)$ is windowed around x_1 by the design the antenna system's directivity.

For a point reflector problem, the values of σ_1 , x_1 and r_1 can be estimated from $s_1(t)$. For a more-realistic SAR problem, the values of many σ_i , x_i and r_i must be estimated from the composite signal

$$s(t) = \sum_i \sigma_i \cos \left\{ \omega t - \frac{4\pi r_i}{\lambda} - \frac{2\pi}{\lambda r_i} (vt - x_i)^2 + \phi_i \right\}. \quad (1.8)$$

Rather remarkably, it turns out to be possible to construct a two-dimensional image by decomposing the one-dimensional signal $s(t)$. Even more remarkable is that the necessary processing was elegantly implemented in the analog optical domain in the 1950s and 60s, well before digital implementation became possible. An exposition of early SAR processing and its implementation with coherent optical methods may be found in Harger [1] and works cited therein.

The point reflector response $s_1(t)$ is a sinusoid with quadratically varying phase, known as a *chirp waveform*. That the received signal is not a pure sinusoid of linearly varying phase is due to the *Doppler effect*. This effect yields an instantaneously-varying frequency with chirp rate

$$\zeta = 2 \frac{\omega v^2}{r_1 c} \quad (1.9)$$

with zero Doppler shift when the SAR is closest to the point reflector, at $x = x_1$. The Doppler effect gives distinct signatures in $s_1(t)$ for any given r_1 and x_1 , and is therefore the critical phenomenon which allows the separation of many point reflectors at coordinates (r_i, x_i) from the received waveform. Neglecting the phase constant ϕ , a baseband-converted receive signal given a point reflector at (r_1, x_1)

may be characterized by

$$r_1(t) = s_{1B}(t) + n(t) \quad (1.10)$$

$$= \sigma_1 \cos \left\{ \frac{2\omega r_1}{c} + \frac{\omega}{r_1 c} (vt - x_1)^2 \right\} + n(t) \quad (1.11)$$

where $n(t)$ is a white noise process. The maximum likelihood estimate of the reflector strength σ_1 , given $x_1 = 0$, r_1 , and $r_1(t)$, is given by the inner product

$$\hat{\sigma}_1 = \int_{-T/2}^{T/2} r_1(t) \varphi(t, r_1) dt \quad (1.12)$$

$$\triangleq \langle r_1(t), \varphi(t, r_1) \rangle \quad (1.13)$$

where the function $\varphi(t, r_1)$ a normalized version of the expected (noiseless) waveform with x_1 taken as zero:

$$\varphi(t, r_1) \triangleq \frac{\cos \left\{ \frac{2\omega r_1}{c} + \frac{\omega}{r_1 c} (vt)^2 \right\}}{\int_{-T/2}^{T/2} \cos^2 \left\{ \frac{2\omega r_1}{c} + \frac{\omega}{r_1 c} (vt)^2 \right\} dt}. \quad (1.14)$$

The inner product in equation 1.13 is a standard matched filter which can be implemented in the digital domain by a FIR filter whose tap weights are a uniform sampling of the function $\varphi(-t, r_1)$. Because the matched filter weights are not a function of x_1 , simply presenting samples of the received signal $r_1(t)$ as input to the filter, will produce a sequence of reflector strength estimates $\hat{\sigma}_1$ valid for successive values of the azimuth variable x_1 . A bank of k such filters, each designed using different range filters $\varphi(-t, r_i)$, $1 \leq i \leq k$, will provide a column of reflector strengths at a constant azimuth. The resulting algorithm is analogous to a stripchart; this mode of SAR is called the *stripmap mode*.

The maximum likelihood assumptions generating $\hat{\sigma}$ are not generally valid for

multiple point reflectors at asynchronous locations. It is a happy coincidence that the underlying Doppler effects produce matched filters which respond strongly to reflectors at the test coordinate (x, r) , but which do not respond in a significant way to signals generated by reflectors far enough away from the test coordinate. This statement is made quantitative by a version of range *ambiguity function*, which in this case is the two-dimensional response of the SAR system to a point reflector situated at a test coordinate (x, r) .

Figure 1.3 shows a typical two-dimensional distance ambiguity function for an X-band radar for a point reflector at $r_1 = 1$ km and $x_1 = 0$. An ideal ambiguity function is a two-dimensional delta function, or *thumbtack*. In practice, the width of the ambiguity function in the range direction (the r variable), and the width in the azimuthal direction (the x variable), are measures of the achievable resolution of the system. The figure shows a resolution on the order of 1 meter in azimuth, and on the order of 10 meters in range. The resolution in azimuth is governed by the autocorrelation function of a chirp waveform; the resolution in range is governed by the cross correlation of two chirp waveforms of differing chirp rate.

This example illustrates the important concept of *pulse compression*. The point reflector is in the beamwidth of the radar for 200 meters of azimuth. The matched filters focus this 200 meter's worth of relatively evenly-spaced chirp energy into an approximation of a delta function with most of its energy distributed within a few meters. Pulse compression of another kind is simultaneously occurring in the range dimension. Because the compression is dependent chiefly upon the characteristics of the radar waveform, the study of more-general transmit waveforms which give better thumbtacks remains a lively research topic to this day.

The system performance of the ambiguity function generally improves with more prominent Doppler effect, that is, with increased radar platform velocity and in-

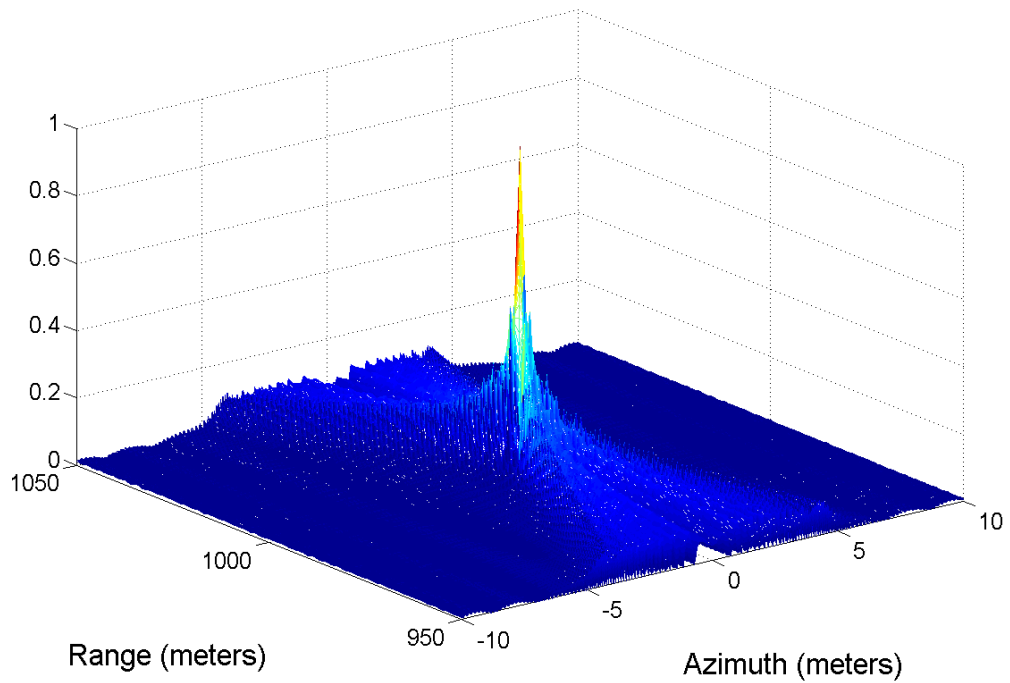


Figure 1.3: Normalized range/azimuth ambiguity function for an X-band (10 GHz) SAR system of velocity 250 meters/second with a point reflector at range $r_1 = 1$ km, azimuth $x_1 = 0$ meters, Doppler bandwidth = ± 10 kHz.

creased transmit frequencies. Improvements in these areas are not always possible, and often resolution performance must be improved when these parameters are decreased for system reasons. To solve these problems, the Doppler phenomenon can be (in effect) synthesized without need for motion at all if the transmit waveform itself is chirped. Many standard SAR algorithms today assume the transmit waveform is a chirp.

1.6 Radar tomography

The maturity of pulse-compression technologies eventually lead to the recognition that the essential measurement in radar approximates a line integral of a reflectivity function over a curve which depends on the geometry of the problem; therefore, radar problems and tomography are fundamentally related.

By 1980, applied mathematicians working in diverse areas, including geophysics and sonar signal processing [26, 27, 28], made significant contributions to the applied theory of generalized Radon transforms. Of particular interest to the present work are Norton's 1980 achievements in sonar processing [27, 28]. Norton developed a physical model for which sonar echoes represented the line integrals, over circles, of a two-dimensional reflectivity function. In this framework, the two dimensional reflectivity function on the upper half (x, y) plane is mapped to a measurement space (x, r) where x is the location of a sonar transponder on the x axis and r is the radius of the line integral's circle with x as the center. Norton developed two inversion methods for this transform. He also developed a closed-form solution giving a function interior to a given circle by its line integrals over circular paths whose centers lie on that circle.

At about the same time as Norton's work, Munson showed in a landmark paper

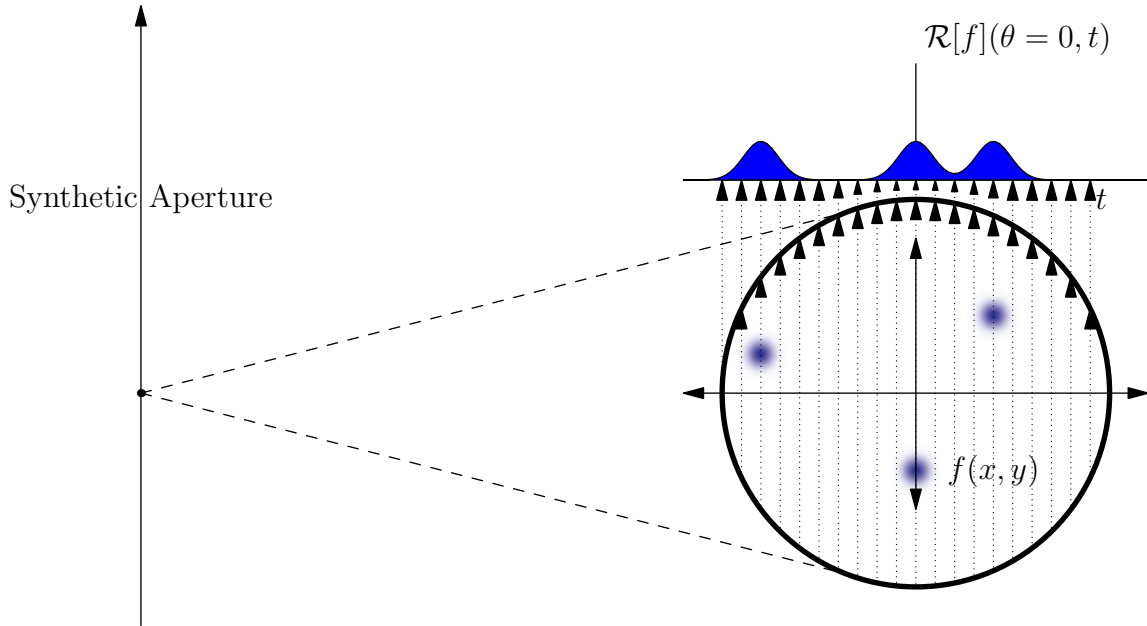


Figure 1.4: Illustration of the connection between spotlight-mode SAR processing and computed tomography. (Figure inspired by J. A. Richards.)

[29] that the reconstruction methods of an important modality of SAR (called the *spotlight mode*) were fundamentally similar to those of computed tomography. Figure 1.4 illustrates this connection. In spotlight-mode SAR, a vehicle flies in a straight line over a two-dimensional reflectivity field. A side-looking, steerable, and directional transmit antenna successively spotlights the same piece of ground as the vehicle moves along its trajectory. A receive antenna on the aircraft records the return signals for each illumination. The trajectory is called the *synthetic aperture*.

Munson and others depended on variations of a generalized *point reflector* model, in which the return signal as a function of time is proportional to the sum of all reflections along contours equidistant from the antennas. In other words, the received waveform at each point on the aperture represents a set of definite line integrals of the 2-D reflectivity function taken along contours generally oriented in the cross-range direction. If the distance between the antennas and the ground of interest

is large, and the directionality of the antennas is sharp, then the received signal at different times represents integrals over different straight lines. Furthermore, the measurements taken at different points on the aperture are line integrals of the scene taken at different angles. With these approximations, the reconstruction of the SAR scene is formally identical to the inversion of the simple Radon transform. By this time, that problem had been efficiently and robustly worked out in the CT arena.

Subsequently, several SAR researchers pointed out that the geometrical approximations used to make SAR look like CT did not have to be made to retain the general idea. Milman [30], Carrera et al [31], Franceschetti et al [32], and Soumekh [33] appear to be the major contributors in this area. All of these researchers worked out variations on the theme of reconstruction by inversion of a generalized Radon transform in which the line integrals are not necessarily over straight lines, but rather over other open conic sections which are determined by the geometry of various SAR modes.

Milman recognized, as explained in his unpublished paper [34], an underlying hyperbolic geometry in the SAR spotlight mode in which the response of a point reflector appears as a hyperbola in the forward transform (that is, “the measurement”) of his model. He used this insight to develop an analytical reconstruction formula similar in form to that used in CT filtered backprojection methods, except that Hankel transforms replace Fourier transforms in various stages of the reconstruction.

Current research in SAR engineering of this type makes further abstractions and generalizations. Redding et al. [2, 35, 36] abstracted the problem into one more purely related to modern integral geometry by completely removing the detail of the transmit waveform and assuming that the radar transmits a delta function. Furthermore, they relaxed the typical side-looking restriction by considering a *down-looking* spherically-radiating antenna. In this case, SAR measurements are modeled

by a *circular Radon transform* given as:

$$\mathcal{R}_c[f](u, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(t - \sqrt{(x - u)^2 + y^2}) dx dy \quad (1.15)$$

which represents the line integrals over all circles of variable radius t centered at location u on the x axis. These measurements are those taken by a monostatic, spherically-radiating SAR platform moving down the x axis at some height above a flat ground. In [2] the authors survey several inversion algorithms for the forward circular transform.

Most recently, using a similar framework also typified by measurements which are line integrals over circles, Yarman [8, 9] demonstrated an inversion algorithm and numerical examples of image reconstruction based on the Funk transform.

1.7 Thesis summary

This work develops and extends a set of signal-processing technologies in support of a simplified-hardware multistatic SAR imaging system.

In Chapter 2, we develop an omnidirectional, multistatic SAR measurement framework in which measurements represent elliptical line integrals of a scene's reflectivity function. The multistatic generalization adds considerable mathematical complexity to the standard monostatic case, even for the simplest multistatic UAV trajectory. Backprojection techniques are thoroughly analyzed in both spatial and frequency domains. We introduce the tomographic method of *approximate inverse*, which produces high quality images for general multistatic trajectories.

Chapter 3 addresses the fundamental and novel question of multistatic UAV trajectory design for high-quality image reconstruction in the presence of noise. The

power of matrix analysis is applied to generate design guidelines for UAV trajectories. Several mathematical bounds are derived and related to physical and geometric aspects of the SAR system. We develop an example multistatic trajectory which greatly improves upon the performance of the standard trajectory. Furthermore, we show that the performance of *random location* trajectories, for large enough problems, might be expected to perform as well as any other measurement strategy.

Chapter 4 addresses options for the simplification of SAR receiver hardware by the design of transmit waveforms giving good pulse compression performance when analyzed by a simple-to-implement Discrete Wavelet Transform filter. The idea of zero-correlation zone waveforms is generalized for the subject sequences to include non-cyclic transmit sequence extensions, providing a novel tradeoff between transmit power and receiver complexity.

Chapter 5 concludes with final remarks and suggestions for future work in this area.

Chapter 2

An elliptical formulation for multistatic omnidirectional SAR

The present work contributes to the state of the art in two areas. First, the previous monostatic frameworks are extended to include general multistatic trajectories. The potential for hardware simplicity of unfocused transmit and receive antennas are shown to extend to the multistatic case. It is shown that an omnidirectional multistatic radar measures the line integrals of a reflectivity function over a set of ellipses, rather than circles, whose parameters depend on the relative location of the transmitter and the receiver over the plane. An integral transform, named the elliptical Radon transform, is proposed to describe the simplest multistatic trajectory: that of a single transmit/receive pair moving in parallel at equal heights over a flat reflective plane.

Second, two methods for recovering the reflectivity function from these measurements are developed. A frequency-domain backprojection algorithm is developed, and compared to simpler algorithms from the monostatic case. The adjoint operator for the forward transform is found to be a type of weighted backprojection. Numerical

versions of these backprojections, which produce approximate images, are compared to one another. Finally, the method of approximate inverse, developed for tomography problems by Louis [37], is adapted to the SAR problem. This method is easily amenable to handle physical deviations from mathematical ideality for a number of important effects. The method is easily adapted to general multistatic trajectories, which are shown to give much more stable results than the simple bistatic trajectories considered here.

The chapter is organized as follows. In section 2.1, we summarize the SAR measurement process in terms of an integral transform of the a two-dimensional reflectivity function. The elliptical Radon transform is introduced to describe the measurements of the simplest bistatic trajectory, and the generalizations needed to describe transforms of a general multistatic trajectory are given. Section 2.2 develops a frequency-domain backprojection method for the elliptical transform, and another backprojection method based on the adjoint of that transform is developed in section 2.3. Section 2.4 contains numerical results for each of the backprojection methods on a simulated Shepp-Logan phantom. In section 2.5, we point out several practical concerns inherent in the preceding results, and solve them by employing the more-general method of approximate inverse; the improved numerical results of this technique are given in section 2.6. The work concludes with final remarks in section 2.7.

2.1 A proposed bistatic SAR geometry

The following geometrical extension of Redding's model [36] treats the more general case of bistatic and multistatic radars over a flat plane of reflectivity $f(x, y)$. For reasons of tractability, the following simplifications made in the development of this

measurement framework. These assumptions are also made in the previous work, and are important enough to enumerate clearly.

1. The widebeam signal radiates spherically from the transmit antennas with zero directivity at all frequencies of interest.
2. Signals propagate in a homogeneous medium at a constant velocity (the speed of light in the medium).
3. The reflections off a point (x, y) in the flat plane are fully characterized by a scalar real reflectivity $f(x, y)$ which is not a function of the angle of incidence.
4. Only primary (single) reflections reach the receive antenna. Secondary reflections due to multipath are assumed to be negligible.
5. The velocities of the UAVs are small enough compared with the speed of light such that the *stop and shoot* approximation can be made, that is, the receivers do not move significantly while receiving a transmit burst.
6. The transmit waveform is either impulsive, or, can be made to look adequately impulsive by modulation and pulse-compression methods.

As Redding and others have shown, most of these assumptions can be relaxed with a variety of impacts in complexity. These assumptions allow a reasonably simple description of SAR measurements in terms of an integral transform of the function $f(x, y)$.

2.1.1 An elliptical Radon transform

Figure 2.1 depicts a transmit and a receive UAV at heights h_1 and h_2 directly over the y axis with y coordinates $\pm d$ from the origin. If the UAVs are at some equal

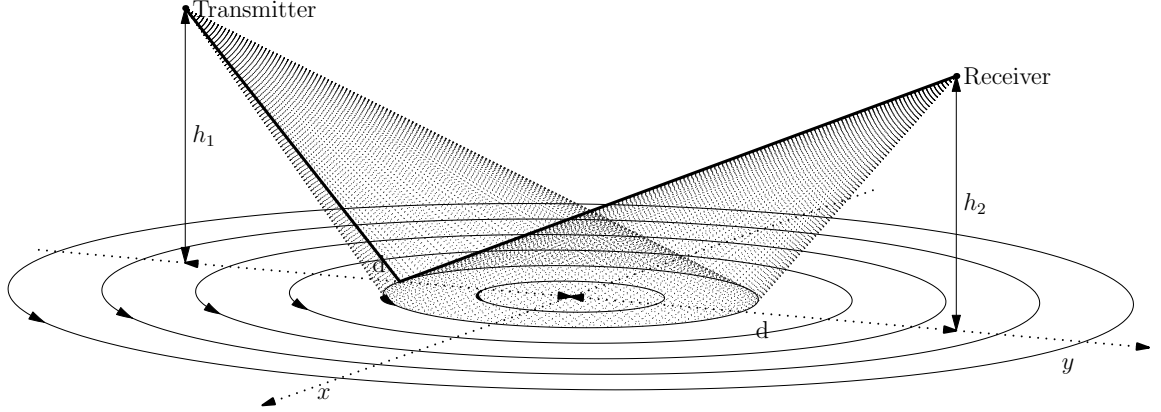


Figure 2.1: Geometrical specification for a single transmit/receive bistatic pair over a flat plane.

height $h = h_1 = h_2$, then a photon moving from the transmitter and reflected to the receiver from a point (x, y) in the plane travels a distance K :

$$K = \sqrt{h^2 + (y + d)^2 + x^2} + \sqrt{h^2 + (y - d)^2 + x^2} \quad (2.1)$$

The signal arriving at the receiver at a time instant $t = K/c$ is that due to all signals which have traveled a distance K . The direct path between the transmitter and receiver is useful for synchronization purposes, but signals from this path give no direct information about the image. The geometry dictates that the signal must have reflected from loci in the xy plane given by a re-arrangement of (2.1):

$$x^2 + y^2 \left(1 - \frac{4d^2}{K^2}\right) = \frac{K^2}{4} - h^2 - d^2 \quad (2.2)$$

which is ellipse centered at the origin for K , h , and d constant. Equation 2.2 may be recast in terms of spatial variables completely in the xy plane by introducing a

new variable $r^2 = K^2/4 - h^2 - d^2$:

$$x^2 + y^2 \frac{r^2 + h^2}{r^2 + h^2 + d^2} = r^2, \quad (2.3)$$

where the semiminor axis lies on the x axis and is of length r , and the semimajor axis lies on the y axis and is of length

$$y_r = r \sqrt{\frac{r^2 + h^2 + d^2}{r^2 + h^2}}. \quad (2.4)$$

The aspect ratio of the ellipse changes smoothly from $\sqrt{1 + (d/h)^2}$ at small r to that of a circle in the limit of large r . Furthermore, if the SAR is monostatic ($d = 0$) then the contours are always circles and the problem degenerates to that considered previously.

The waveform seen at the receiver undergoes filtering and pulse compression as a function $f(t)$. This function can be converted to the spatial domain $f(r)$ by a simple resampling using the relationship

$$t = \frac{2}{c} \sqrt{r^2 + d^2 + h^2} \quad (2.5)$$

where the origin of the t dimension is the moment of transmission of the effective impulse. In numerical versions of image reconstruction (discussed in a later section), it will be convenient to uniformly sample the function $f(r)$. Because the relationship between r and t is not linear, such a sampling corresponds to a nonuniform sampling of the function $f(t)$.

The received signal as t increases from the instant of transmission is due to reflections from an expanding set of ellipses. The ellipses are centered in the xy plane halfway between the UAVs, whose major axis lies on the straight line connecting the

projection of the UAVs' location in the xy plane; the ellipses' eccentricity changes as a function of time and radius. The amplitude of the received signal at the variable r corresponding to the time t is proportional to the sum of the reflectivity function around the corresponding ellipse, that is, a line integral of $f(x, y)$.

The signal is also proportional to some function $\beta(K)$ which represents a *path loss* of the signal as it traverses the distance K . As Norton realized [28], the received signal at any point in time is that from a constant distance K and therefore may be said to have a path loss $\beta(K)$ at that point. Therefore the path loss can be removed from the inversion problem by dividing $\beta(K)$ out of the radar measurement. Under this model, the path loss effect does not play a major role in the inversion process; however, clearly $\beta(K)$ cannot be ignored if the effects of additive noise in the measurement are to be treated properly.

The full measurement made by a single transmission burst is a set of line integrals of the reflectivity $f(x, y)$ over an expanding set of ellipses. If the ellipses are parameterized in local polar-like coordinates such that

$$x = r \cos \theta \tag{2.6}$$

$$y = r \sqrt{\frac{r^2 + h^2 + d^2}{r^2 + h^2}} \sin \theta, \tag{2.7}$$

then the single-burst, bistatic measurement with center at the origin of the xy plane, and at the time corresponding to the radius r , is

$$\begin{aligned} F(r) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x^2 + y^2 \frac{r^2 + h^2}{r^2 + h^2 + d^2} - r^2) dx dy \\ &= r \int_0^{2\pi} f(r \cos \theta, r \sqrt{\frac{r^2 + h^2 + d^2}{r^2 + h^2}} \sin \theta) \sqrt{1 + \frac{d^2}{r^2 + h^2} \cos^2 \theta} d\theta, \end{aligned} \tag{2.8}$$

where the differential length dl of the line integral is written in terms of the integra-

tion variable $d\theta$ as:

$$dl = r \sqrt{1 + \frac{d^2}{r^2 + h^2} \cos^2 \theta} d\theta. \quad (2.9)$$

An appropriate generalized forward Radon transform which represents a set of measurements of this type can be defined. The simplest such transform corresponds to a bistatic pair moving in slow time parallel to the x axis with spatial coordinates $(x, y, z) = (vt, \pm d, h)$. An elliptical Radon transform, a generalization of Redding's circular Radon transform, may be defined as

$$\mathcal{R}_e[f](x, r) \triangleq r \int_0^{2\pi} f(r \cos \theta + x, r \sqrt{\frac{r^2 + h^2 + d^2}{r^2 + h^2}} \sin \theta) \sqrt{1 + \frac{d^2}{r^2 + h^2} \cos^2 \theta} d\theta, \quad (2.10)$$

where $\mathcal{R}_e[f]$ is a function of x and r . Methods to recover the function $f(x, y)$ from $\mathcal{R}_e[f]$ are the subject of the next section.

If the value of d approaches zero, the SAR becomes monostatic, and the transform reduces to the considerably-simpler circular Radon transform:

$$\lim_{d \rightarrow 0} \mathcal{R}_e[f] = r \int_0^{2\pi} f(r \cos \theta + x, r \sin \theta) d\theta = \mathcal{R}_c[f]. \quad (2.11)$$

In common with the circular Radon transform and the Funk transform, inversion of \mathcal{R}_e can only hope to recover the even part of $f(x, y)$ with respect to y . The symmetry of the integration contours dictates that all functions $f(x, y)$ which are odd with respect to y will be members of the null space of \mathcal{R}_e .

Further insight into the forward transform \mathcal{R}_e may be gained by studying the measurement of a point reflector situated at coordinates (x_0, y_0) . In this case, $f(x, y) = \delta(x - x_0, y - y_0)$, and $\mathcal{R}_e[f]$ is a line delta function, which is zero at all x , and

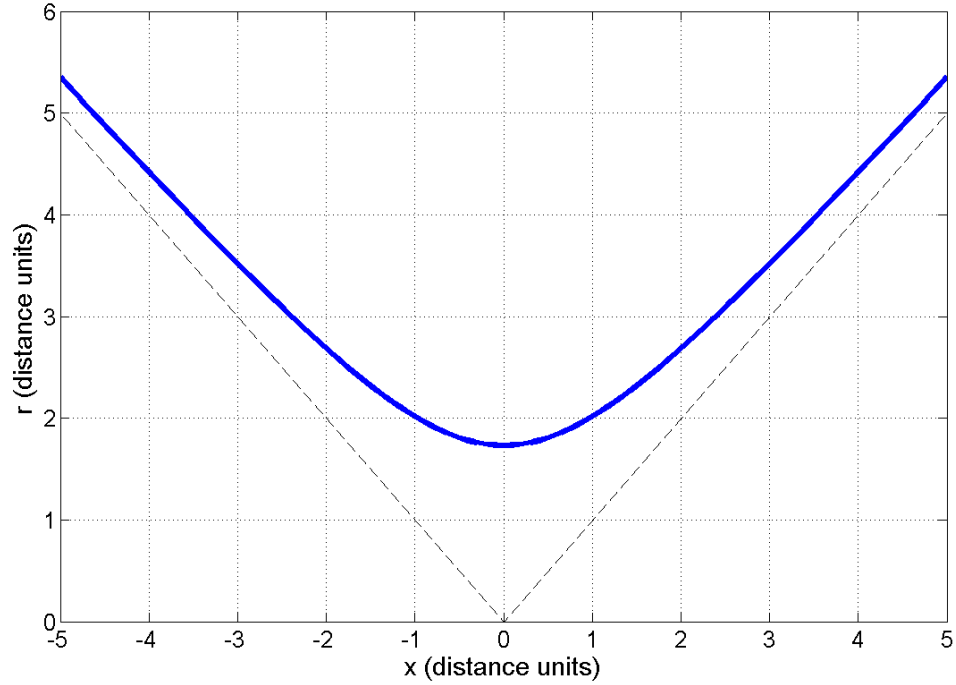


Figure 2.2: The point reflector response $\mathcal{R}_e[\delta(x_0, y_0)]$ for $x_0 = 0$, $y_0 = 2$, $h = 3$, and $d = 2$.

at all $r > 0$, except on the curve

$$2r^2 = (x - x_0)^2 + y_0^2 - h^2 - d^2 + \sqrt{((x - x_0)^2 + y_0^2 + h^2 + d^2)^2 - 4d^2y^2}. \quad (2.12)$$

Figure 2.2 shows this curve for some given values of its parameters. The curve is similar to simple hyperbola, and will approach the asymptote $r = |x - x_0|$ when x is far from x_0 . Therefore (in common with many transforms), $\mathcal{R}_e[f]$ of some compactly-supported function f will generally have infinite support in the (x, r) plane. Viewed as a two-dimensional spatial impulse response, $\mathcal{R}_e[\delta(x_0, y_0)]$ is clearly linear but it is not shift-invariant. More specifically, the shape of the response does not vary with x_0 ; the shape does vary with y_0 . Therefore, a simple inversion by two-dimensional

deconvolution is not possible in this form.

In the previously-considered monostatic case ($d = 0$), the impulse response simplifies to a line delta function along the simpler curve

$$r^2 = (x - x_0)^2 + y_0^2, \quad (2.13)$$

which is exactly a hyperbola in the (x, r) plane. Norton [28] first attempted solution of this inversion problem in 1980 by warping the y dimension by the transformation $\zeta = y^2$. In this new space, the impulse response is the shift-invariant parabola $r^2 = (x - x_0)^2 + \zeta_0$, for which a straightforward inversion based on deconvolution may be possible (however, Redding and Newsam dispute the validity of Norton's argument in [36]). In any event, there is no such simple transformation which suffices to transform the more general case, considered in (2.12), to be shift invariant.

2.1.2 General integral transforms

Clearly, much-more-general integral transforms may be defined which represent integrals over families of ellipses whose centers are generated by the curved trajectories of many UAVs. A more-general case is depicted in Figure 2.3, in which the axes of the measurement ellipses are not necessarily parallel with the reference axes. Generally, any SAR measurement process simply generates some sampling of the line integrals around all possible ellipses in the plane. Whether a given sampling strategy is achieved by bistatic pairs over a long time, or multistatic groups over a shorter time, is not a fundamental mathematical distinction for the reconstruction of a static scene. Some sampling strategies will be more efficient than others.

Finally, the measurement space can be further generalized by allowing the trajectories themselves to be three dimensional, that is, by permitting the UAVs to

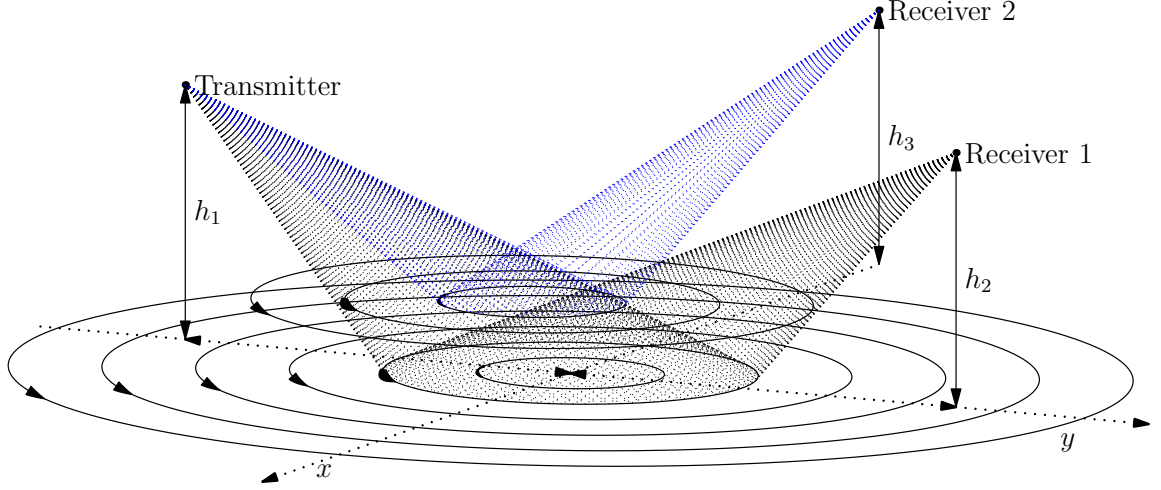


Figure 2.3: Geometrical depiction of a simple multistatic configuration over a flat plane.

wander in the z direction in addition to the x and y directions. A straightforward generalization of equation 2.2 for non-equal UAV heights h_1 and h_2 yields

$$x^2 + \left(1 - \frac{4d^2}{K^2}\right)\left(y - \frac{d(h_1^2 - h_2^2)}{K^2 - 4d^2}\right)^2 = \frac{K^2}{4} - d^2 - \frac{h_1^2 + h_2^2}{2} + \frac{(h_1^2 - h_2^2)^2}{4(K^2 - 4d^2)}. \quad (2.14)$$

Setting the right hand side to r^2 leads to the re-parametrization of the ellipses:

$$x^2 + \frac{\hat{r}^2}{\hat{r}^2 d^2} \left(y - \frac{4d(h_1^2 - h_2^2)}{\hat{r}^2}\right)^2 = r^2 \quad (2.15)$$

with $\hat{r}^2 = \frac{1}{4}\{h_1^2 + h_2^2 + 2r^2 + 2\sqrt{(r^2 + h_1^2)(r^2 + h_2^2)}\}$. These relations can lead to a still-more-general set of Radon forward transforms, based on more-general vehicle trajectories. The equation describes ellipses of integration; however, with unequal heights, the center of the ellipses also changes as a function of time or distance. This behavior adds additional challenges to the analysis, but also suggests interesting three dimensional trajectories. Downward-spiraling, helical paths might be of high interest for non-powered glider vehicle. The diversity in ellipse centers provided by

differential heights also suggests easily-implemented (if short-lived) flight patterns in which all vehicles fall straight down. Radar measurements using any such multistatic trajectory can be described under this framework, simply by rotating the local x and y axis defined by a given transmitter/receiver instantaneous geometry, to a common reference frame.

2.2 Backprojection operators for \mathcal{R}_e

Frequency-domain backprojection methods very often form the basis for the numerical inversion of many types of Radon transforms. Both the ordinary x-ray Radon transform and the circular Radon transform can be inverted by filtering the transform's adjoint backprojection operator [23, 36]. In both cases, the backprojection and the filter can be efficiently computed in the two-dimensional frequency domain.

In this section, a frequency-domain backprojection operator for $\mathcal{R}_e[f]$ is developed. Backprojection results for the simpler circular Radon transform indicate that the interplay of the Fourier and Hankel transforms are important for this type of problem. The one-dimensional Hankel transform appears because it may be viewed as a two-dimensional Fourier transform of a two-dimensional function which has circular symmetry about the origin. As shall be seen, a new and more-complicated Hankel-like transform, for two-dimensional functions possessing a kind of *elliptical* symmetry, results from this investigation.

Using the standard backprojection intuition, a backprojected function $f_B(x, y)$ may be generated from $\mathcal{R}_e[f] \triangleq g(x, r)$ by

$$f_B(x, y) = \int_{-\infty}^{\infty} \frac{g(u, r_u)}{L(r)} du \quad (2.16)$$

where r_u is shorthand for the function of x and y which is the solution of (2.3) for r , with x shifted by the integration variable u :

$$r_u^2 = \frac{1}{2} \left\{ (x - u)^2 + y^2 - h^2 - d^2 + \sqrt{((x - u)^2 + y^2 + h^2 + d^2)^2 - 4d^2y^2} \right\}, \quad (2.17)$$

and $L(r)$ is some normalizing function of the forward transform variable r_u . For every ellipse center u , there exists a single ellipse passing through the point (x, y) . The backprojected contribution to f_B at (x, y) is that ellipse's line integral of $f(x, y)$, which is $g(u, r_u)$, normalized by the length of that ellipse. This is the integrand of (2.16). The integral sums up the contributions over all ellipse centers u .

Several varieties of backprojection exist, depending on various reasons for weighting the measurement $g(u, r_u)$. In previous work on the circular Radon transform[36], the function $L(r)$ was chosen to be unity because then the operation is identical to the adjoint operation on the forward transform. (The question of the adjoint operator for \mathcal{R}_e is taken up in the next section). In another version, $L(r)$ might be taken to be the length of the sampling ellipse, that is,

$$L(r) = r \int_0^{2\pi} \sqrt{1 + \frac{d^2}{r^2 + h^2} \cos^2 \theta} d\theta. \quad (2.18)$$

When $d \neq 0$, $L(r)$ is an elliptic special function of r . The use of this version of L constrains the contribution to the backprojection, of each ellipse corresponding to $g(u, r)$, to have the same line integral as that of the original function over the same ellipse. The results of this section, however, are valid for any choice of L .

The integration of (2.16) can be carried out numerically using quadrature algorithms and re-interpolations of a given sampled function $g(x_n, r_m)$. But (2.16) admits no obvious “fast” algorithm for the integration. However, as shall be seen

next, transformation of f_B to the two-dimensional frequency domain *does* eventually lead to standard Fourier and Hankel transforms of re-scalings of the function $g(x, r)$, for which fast algorithms exist.

2.2.1 Frequency-domain backprojection

A Fourier analysis of the backprojection problem begins by developing the two-dimensional Fourier transform of an elliptical disk centered at the origin, as a useful intermediate result. Bracewell solved the simpler problem of a circular disk in [25]; here, his derivation is generalized for the elliptical-disk problem.

If a function $f_D(x, y)$ is unity interior to the ellipse defined by (2.3), and zero exterior to it, then the two-dimensional Fourier transform of $f_D(x, y)$ is

$$F_D(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_D(x, y) e^{-2\pi j(ux+vy)} dx dy \quad (2.19)$$

$$= \int_{-r}^r e^{-2\pi jux} \int_{-\alpha\sqrt{r^2-x^2}}^{\alpha\sqrt{r^2-x^2}} e^{-2\pi v y} dx dy \quad (2.20)$$

$$= \frac{2}{\pi v} \int_0^r \cos(2\pi ux) \sin(2\pi \alpha v \sqrt{r^2 - x^2}) dx \quad (2.21)$$

where for shorthand the parameter

$$\alpha^2 \equiv \frac{r^2 + h^2 + d^2}{r^2 + h^2} \quad (2.22)$$

is useful.

The definite integral in (2.19) has already been worked out (see GR 3.711 [38]).

The solution is

$$F_D(u, v) = \frac{r}{\sqrt{u^2 + \alpha^2 v^2}} J_1(2\pi r \sqrt{u^2 + \alpha^2 v^2}) \quad (2.23)$$

where J_1 is a Bessel function of the first order. By (2.23), $F_D(u, v)$ also shows simple

elliptical symmetry in the Fourier domain, with foreshortening in the v direction by the parameter $\alpha(r)$.

Equation 2.23 can be used to determine the two-dimensional Fourier transform of a more complicated function $f_r(x, y)$ with the elliptical symmetry, about the origin, of (2.3). In this case, $f_r(x, y)$ is determined everywhere on the plane by its values on the positive x axis according to the rule:

$$f_r(x, y) = f_r(r \cos \theta, r\alpha \sin \theta) = f_r(r) \quad (2.24)$$

for all $r \geq 0$ and all θ and some real function $f(r)$. Note that in such a case,

$$\mathcal{R}_e[f_r](x, r)|_{x=0} = f_r(r) \int_0^{2\pi} r \sqrt{1 + \frac{d^2}{r^2 + h^2} \cos^2 \theta} d\theta = f_r(r)L(r) \quad (2.25)$$

where $L(r)$ is the length of the ellipse of integration, and where the function $f_r(r)$ may be taken outside the integral because it does not change along the line of integration.

If $f_r(r)$ is differentiable and vanishing at large r , then the Fourier transform of the elliptically symmetric function $f_r(x, y)$ can be constructed by accumulating the contributions due to elliptical disks of differential amplitude $-df_r(r)$ and x -axis radius r :

$$\begin{aligned} F_r(u, v) &= - \int_0^\infty \frac{df_r(r)}{dr} F_D(u, v) dr \\ &= -f_r(r)F_D(r)|_0^\infty + \int_0^\infty f(r)F'_D(r) dr \\ &= \int_0^\infty f_r(r)F'_D(r) dr \\ &= \int_0^\infty \frac{g_r(x, r)|_{x=0}}{L(r)} F'_D(r) dr \end{aligned} \quad (2.26)$$

where the first step is due to integration by parts and the second step follows from $J_1(0) = 0$ and $f(\infty) = 0$. Through use of the Bessel identities $2J_1(x) = x(J_0(x) +$

$J_2(x)$) and $2J_1'(x) = J_0(x) - J_2(x)$ [39], and changing to the frequency-domain polar coordinates (ρ, ϕ) , where $\rho = \sqrt{u^2 + v^2}$ and $\tan \phi = v/u$, F_r may be written as:

$$\begin{aligned}
F_r(\rho, \phi) = & \int_0^\infty \frac{g_r(x, r)|_{x=0}}{\frac{1}{2\pi}L(r)} r J_0(2\pi r \rho \sqrt{1 + \frac{d^2}{r^2 + h^2} \sin^2 \phi}) dr \\
& + \int_0^\infty \frac{g_r(x, r)|_{x=0}}{\frac{1}{2\pi}L(r)} \frac{d^2 r^2 \sin^2 \phi}{(r^2 + h^2)(d^2 \sin^2 \phi + r^2 + h^2)} r J_2(2\pi r \rho \sqrt{1 + \frac{d^2}{r^2 + h^2} \sin^2 \phi}) dr.
\end{aligned} \tag{2.27}$$

By construction, F_r is an integral transformation of a measurement vector $g_r(x, r)|_{x=0}$, giving the two-dimensional Fourier transform of some elliptically symmetric $f_r(x, y)$. Now, consider the subset of general functions $f(x, y)$ such that $g(x, r)|_{x=0} = f(r)$ for some given $f(r)$. Each member of this subset, including that unique $f(x, y) = f_r(x, y)$ which actually *is* elliptically symmetric, will have a common elliptically-symmetric backprojection which is equal to $f_r(x, y)$. Therefore, the inverse Fourier transform of F_r , $\mathcal{F}^{-1}F_r = f_r(x, y)$, is the backprojection of the general measurement function $g(x, r)|_{x=0}$ into the spatial domain. The subscripts r will be dropped henceforth, to respect this interpretation of (2.27), and the backprojected approximation of the general $f(x, y)$ from $g(x, r)|_{x=0}$ will be denoted as $f_{B0}(x, y) = \mathcal{F}^{-1}F_{B0}$.

The common argument to the Bessel functions of (2.27) suggests a transformation of the variable r to a new variable R using the relations

$$R \triangleq r \sqrt{1 + \frac{d^2}{r^2 + h^2} \sin^2 \phi} \tag{2.28}$$

$$\frac{dR}{dr} = \frac{(r^2 + h^2)^2 + d^2 h^2 \sin^2 \phi}{(r^2 + h^2)^2 \sqrt{1 + \frac{d^2}{r^2 + h^2} \sin^2 \phi}} \tag{2.29}$$

$$2r^2 = R_-^2 + \sqrt{R_-^4 + 4h^2 R^2} \tag{2.30}$$

$$R_h^2 \triangleq r^2 + h^2 = \frac{1}{2}R_+^2 + \frac{1}{2}\sqrt{R_-^4 + 4h^2 R^2} \tag{2.31}$$

$$R_{\pm}^2 \triangleq R^2 \pm h^2 - d^2 \sin^2 \phi \quad (2.32)$$

leading to the simpler-looking forms

$$F_{B0}(\rho, \phi) = \int_0^{\infty} \mathfrak{g}_0(x, R)|_{x=0} J_0(2\pi\rho R) R dR + \int_0^{\infty} \mathfrak{g}_2(x, R)|_{x=0} J_2(2\pi\rho R) R dR \quad (2.33)$$

$$= \mathfrak{g}_0^{(\mathcal{I}, \mathcal{H}_0)}(x, \rho; \phi)|_{x=0} + \mathfrak{g}_2^{(\mathcal{I}, \mathcal{H}_2)}(x, \rho; \phi)|_{x=0}, \quad (2.34)$$

where $\mathfrak{g}_0(x, R)$ and $\mathfrak{g}_2(x, R)$ are the resampled and rescaled versions of $g_r(x, r)$ derived by changing the variable from r to R :

$$\mathfrak{g}_0(x, R) \triangleq \frac{g(x, R)}{\frac{1}{2\pi}L(R)} \left\{ \frac{R_h^4}{R_h^4 + d^2 h^2 \sin^2 \phi} \right\} \triangleq \frac{g(x, R)}{\frac{1}{2\pi}L(R)} W_0(R) \quad (2.35)$$

and

$$\mathfrak{g}_2(x, R) \triangleq \frac{g(x, R)}{\frac{1}{2\pi}L(R)} \left\{ \frac{d^2 \sin^2 \phi R_h^2 (R_h^2 - h^2)}{(R_h^4 + d^2 h^2 \sin^2 \phi)(R_h^2 + d^2 \sin^2 \phi)} \right\} \triangleq \frac{g(x, R)}{\frac{1}{2\pi}L(R)} W_2(R). \quad (2.36)$$

2.2.2 Discussion

The first term of the right-hand side of (2.33) is identical to the one-dimensional, order-zero Hankel transform of \mathfrak{g}_0 with respect to its second variable, R ; similarly, the second term of (2.33) is the order-two Hankel transform of \mathfrak{g}_0 with respect to R . In the transform notation $\mathfrak{g}_i^{(\mathcal{I}, \mathcal{H}_i)}(x, \rho; \phi)$, the \mathcal{I} in the superscript means that \mathfrak{g}_i undergoes the identity transform, or remains intact, in its first variable x ; the superscript \mathcal{H}_i means that \mathfrak{g}_i is i^{th} -order Hankel transformed in its second variable R . The $\rho; \phi$ denotes that the Hankel transform variable is ρ , with a reminder that the transform has parameter ϕ , because (in this case), R is a function of ϕ .

While the scale changes of $g(x, R)|_{x=0}$ in equations 2.35 and 2.36, shown in braces and named W_i , may appear complex in their detail, in essence they are simply transfer functions. The transfer functions give a kind of “frequency response” which emphasizes or deemphasizes the forward transform depending on the value of the transform variable r . Figure 2.4 shows the shape of these functions for particular values of the parameters d and h , and for various values of the transform variable ϕ . The transfer function for \mathbf{g}_0 is an allpass-type function which can attenuate $g(x, r)$ slightly at small values of r . The amount of attenuation depends on the variable ϕ . The transfer function for \mathbf{g}_2 is a bandpass-type function, making \mathbf{g}_2 responsive only to that part of $g(x, r)$ near (in this case) $r = 3$ distance units. Furthermore, the function \mathbf{g}_2 has negligible effect near the u axis of the function $F_{B0}(u, v)$, and maximum effect near the v axis of the $F_{B0}(u, v)$.

It is appropriate to digress briefly in order to explain away the apparent singularity arising from the normalizing length-of-the-ellipse factor $L(r)$, which commonly appears in the denominator of many of the above equations, and which clearly goes to zero when r or R goes to zero. The denominator $L(r)$ always appears in a ratio with the function $g(x, r)$, which for infinitesimal r can be written as

$$\lim_{r \rightarrow 0} \frac{g(x, r)}{L(r)} = \lim_{r \rightarrow 0} \frac{r \int_0^{2\pi} f(x + r \cos \theta, r \sin \theta) \sqrt{1 + \frac{d^2}{r^2+h^2} \cos^2 \theta} d\theta}{r \int_0^{2\pi} \sqrt{1 + \frac{d^2}{r^2+h^2} \cos^2 \theta} d\theta} \quad (2.37)$$

$$= f(x, 0) \lim_{r \rightarrow 0} \frac{\int_0^{2\pi} \sqrt{1 + \frac{d^2}{r^2+h^2} \cos^2 \theta} d\theta}{\int_0^{2\pi} \sqrt{1 + \frac{d^2}{r^2+h^2} \cos^2 \theta} d\theta} \quad (2.38)$$

$$= f(x, 0) \quad (2.39)$$

so that no singularity exists. Because $g(x, r)$ itself always appears with the $L(r)$ normalization factor in the backprojection equations, it is natural to suggest that

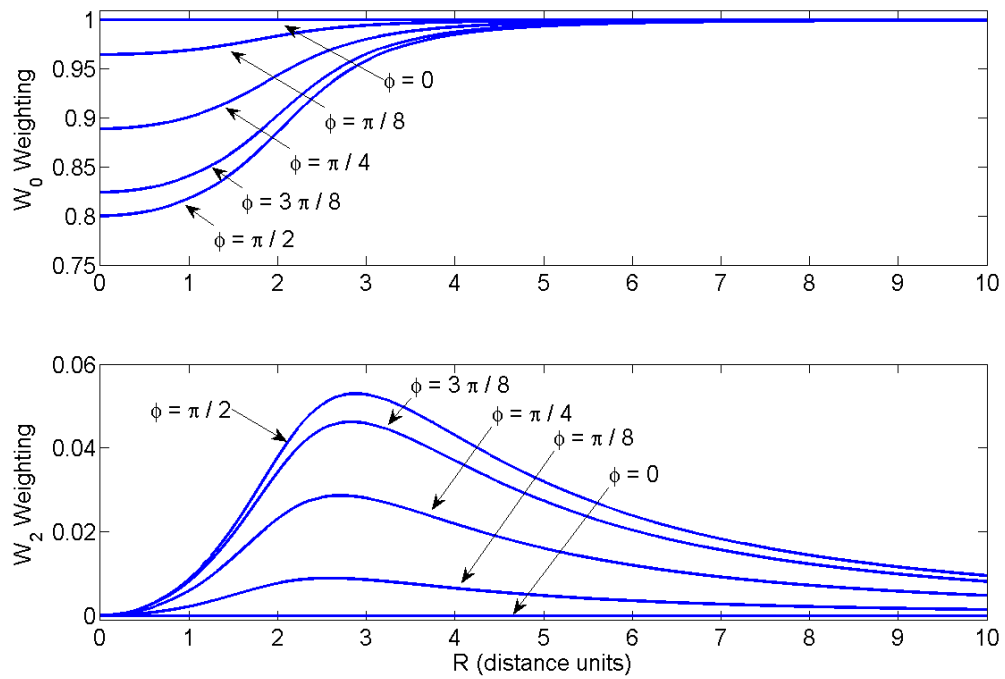


Figure 2.4: The transfer functions W_0 and W_2 versus R using $d = 1$, $h = 2$, and various values of ϕ .

the divisor $L(r)$ simply be included in the definition of the forward transform \mathcal{R}_e . It is easy to see that the apparent singularity is avoided altogether by deleting the initial factor of r in (2.10), as is done by some authors for the circular Radon transform. However, in the present work, the meaning of the original \mathcal{R}_e as something directly related to a physical measurement is retained.

Finally, the total backprojection F_B of the measurement function $g(x, r)$ is the sum of the contributions over all x , not just the contribution at $x = 0$. The contribution at a general x is the F_{B0} of (2.33), but the restriction that the function $g(x, r)$ be evaluated at $x = 0$ is dropped. In addition, because that F_{B0} corresponds to a backprojection centered at the origin, the resulting F_{B0} must be shifted such that its inverse transform is centered at $(x, 0)$, by the standard Fourier shift operator $e^{-j2\pi ux}$. Therefore, using $u = \rho \cos \phi$, the total backprojection is

$$F_B(u, v) = \int_{-\infty}^{\infty} e^{-j2\pi x \rho \cos \phi} (\mathfrak{g}_0^{(I, \mathcal{H}_0)} + \mathfrak{g}_2^{(I, \mathcal{H}_2)}) dx. \quad (2.40)$$

Noting that the integration over all x phases is identical to the one-dimensional Fourier transform with respect to the first variable, x , the result can be written remarkably simply as

$$f_B^{(\mathcal{F}, \mathcal{F})}(\rho, \phi) = \mathfrak{g}_0^{(\mathcal{F}, \mathcal{H}_0)}(\rho \cos \phi, \rho; \phi) + \mathfrak{g}_2^{(\mathcal{F}, \mathcal{H}_2)}(\rho \cos \phi, \rho; \phi) \quad (2.41)$$

where an \mathcal{F} in the function superscript indicates the Fourier transform with respect to the first or second function variable, depending on its position. In conventional notation, each term in the result is of the form

$$\mathfrak{g}_i^{(\mathcal{F}, \mathcal{H}_i)} = \int_{-\infty}^{\infty} \int_0^{\infty} e^{-j2\pi x \rho \cos \phi} R J_i(2\pi \rho R) \left\{ \frac{g(x, R) W_i(R)}{\frac{1}{2\pi} L(R)} \right\} dx dR \quad (2.42)$$

where the portion in braces is implicitly a function of the outer transform variable ϕ . One may efficiently evaluate $f_B^{(\mathcal{F},\mathcal{F})}(\rho, \phi)$ along lines of constant ϕ radiating from the origin. Then, all integrals are the standard one-dimensional Fourier and Hankel transforms.

In the monostatic case ($d = 0$) with circular line integrals, then $r = R$, $W_2 = 0$, $L(r) = 2\pi r$, $W_0 = 1$, and $\mathbf{g}_0 = g(x, r)/r$, giving the result :

$$f_B^{(\mathcal{F},\mathcal{F})}(\rho, \phi)|_{d=0} = \int_{-\infty}^{\infty} \int_0^{\infty} e^{-j2\pi x\rho \cos \phi} J_0(2\pi\rho r) \left\{ \frac{g(x, r)}{r} \right\} dx dr \quad (2.43)$$

$$= \mathbf{g}_0^{(\mathcal{F},\mathcal{H}_0)}(\rho \cos \phi, \rho) \quad (2.44)$$

which is equivalent to the backprojection result found in [36].

The backprojection methods developed in this section are based on standard, but somewhat heuristic, concepts. We now develop another backprojection method based on the adjoint operator of the forward transform.

2.3 The adjoint of \mathcal{R}_e

Simple filtering of the adjoint of the circular Radon transform operator (and that of the standard Radon transform) forms the basis for an inversion technique for those transforms. The purpose of this section is to develop the adjoint of the elliptical transform \mathcal{R}_e .

Following the notation of Redding and Newsam, the inner product on \mathbb{R}^2 is denoted by $\langle \cdot, \cdot \rangle$, and the inner product on $\mathbb{R} \times [0, \infty]$ is denoted by $\langle \cdot, \cdot \rangle_r$. The

adjoint problem is to seek the operator $\mathcal{R}_e^\dagger[g]$ such that $\langle \mathcal{R}_e[f], g \rangle_r = \langle f, \mathcal{R}_e^\dagger[g] \rangle$:

$$\langle \mathcal{R}_e[f], g \rangle_r = \int_{-\infty}^{\infty} du \int_0^{\infty} g(u, r) dr \left\{ r \int_0^{2\pi} f(u + r \cos \theta, \sqrt{\frac{r^2 + h^2 + d^2}{r^2 + h^2}} r \sin \theta) \dots \right. \\ \left. \sqrt{1 + \frac{d^2}{r^2 + h^2}} \cos^2 \theta d\theta \right\} \quad (2.45)$$

The integral over r and θ resembles an area integral over polar-like coordinates, related to the target x and y Cartesian coordinates by

$$x = r \cos \theta, \quad (2.46)$$

$$y = \sqrt{\frac{r^2 + h^2 + d^2}{r^2 + h^2}} r \sin \theta. \quad (2.47)$$

The Jacobian of the transformation from Cartesian to such polar coordinates is

$$J_{xy}(r, \theta) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = r \frac{(r^2 + h^2)^2 + d^2(r^2 \cos^2 \theta + h^2)}{(r^2 + h^2)^{3/2} \sqrt{r^2 + h^2 + d^2}}. \quad (2.48)$$

Recalling that the inverse transformation, from polar to Cartesian, has a Jacobian equal to the reciprocal of $J_{xy}(r, \theta)$, a partial inverse transformation yields

$$\langle \mathcal{R}_e[f], g \rangle_r = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x + u, y) \frac{r}{J_{xy}} \sqrt{1 + \frac{d^2}{r^2 + h^2}} \cos^2 \theta g(u, r) dx dy \quad (2.49)$$

$$= \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x + u, y) W_A(x, r) g(u, r) dx dy \quad (2.50)$$

with a transfer function

$$W_A(x, r) = \frac{(r^2 + h^2) \sqrt{(r^2 + h^2 + d^2(x/r)^2)(r^2 + h^2 + d^2)}}{(r^2 + h^2)^2 + d^2(x^2 + h^2)}. \quad (2.51)$$

Transforming $x + u$ to x and re-arranging the order of the integration gives

$$\langle \mathcal{R}_e[f], g \rangle_r = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \left\{ \int_{-\infty}^{\infty} W_A(x - u, r_u) g(u, r_u) du \right\} dx dy \quad (2.52)$$

$$\triangleq \langle f, \mathcal{R}_e^\dagger[g] \rangle, \quad (2.53)$$

where r_u is as defined previously in (2.17), and where the adjoint operator is recognizable as the weighted backprojection

$$\mathcal{R}_e^\dagger[g(x, y)] = \int_{-\infty}^{\infty} W_A(x - u, r_u) g(u, r_u) du. \quad (2.54)$$

The objects W_A and r_u have not been fully transformed into functions of x and y , for two reasons. First, the transformation in (2.17) produces complicated and unwieldy expressions in x and y , which do not easily produce additional insight. Second, the variable r_u must be used as the second argument to the function $g(u, r_u)$; that value can be re-used directly in the W_A expression when necessary.

The adjoint operator \mathcal{R}_e^\dagger is a backprojection because it integrates the value of the measurement over its ellipse (that is, the values of x and y for which r_u is constant). The magnitude the backprojection is weighted by a non-separable function of the semi-minor axis of the ellipse, r_u , and square of the local abscissa, $(x - u)^2$. These weighting variables are coupled in such a way that the weighting cannot be expressed purely in terms of the transform variables, and in particular, the weighting cannot be expressed as the function of r , $L(r)$. Therefore, the frequency-domain backprojection

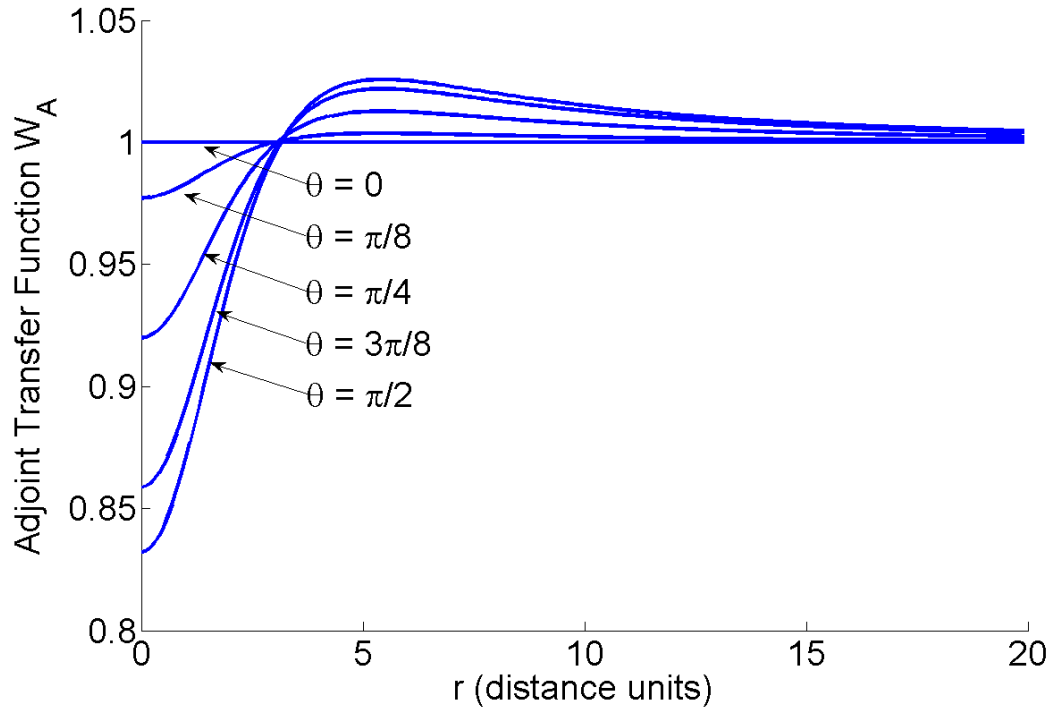


Figure 2.5: The transfer function W_A versus r using $d = 2$, $h = 3$, and various values of $\cos \theta = x/r$.

method developed in the previous section cannot be used for the adjoint.

Recalling that x cannot exceed r in magnitude, because the ratio $x/r = \cos \theta$, the limiting values of the transfer function $W_A(x, r)$ are:

$$\lim_{r \rightarrow \infty} W_A(x, r) = 1, \quad (2.55)$$

$$\sqrt{\frac{h^2}{h^2 + d^2}} \leq \lim_{r \rightarrow 0} W_A(x, r) = \sqrt{\frac{h^2 + d^2(x/r)^2}{h^2 + d^2}} \leq 1. \quad (2.56)$$

Therefore, W_A is an allpass function which deemphasizes the backprojection when the measurement ellipses are relatively small *and* the adjoint variable y is in the upper part of the ellipse. Figure 2.5 shows this behavior for various values of the ratio of x to r .

If $d = 0$, then $W_A(x, r) = 1$, $r_u = \sqrt{(x - u)^2 + y^2}$, and the adjoint $\mathcal{R}_e^\dagger[g]$ reduces to the considerably simpler adjoint for the circular Radon transform found in [36]:

$$\mathcal{R}_c^\dagger[g(x, y)] = \int_{-\infty}^{\infty} g(u, \sqrt{(x - u)^2 + y^2}) du. \quad (2.57)$$

2.4 Numerical backprojection results

The foregoing purely-formulaic results suggest that a numerical comparison of the various backprojection strategies is not inappropriate.

The backprojection problem may be discretized in the following way. The image is regarded as an $N \times N$ rectangular pixel grid, with the measurement value scaled by the length of the measurement-ellipse segment which passes through each pixel. The simplest backprojection method, that of (2.16) using $L(r) = 1$, is then simply given by

$$\hat{f}_{1j} = \sum_{i=1}^m A_{ij} r_i, \quad (2.58)$$

where \hat{f}_j is the estimate of the j^{th} pixel, A_{ij} is a matrix element describing the length of the i^{th} ellipse which is contained within the j^{th} pixel, and r_i is the measurement value associated with the i^{th} ellipse. Similarly, if the measurements are normalized by the ellipse length L_i , then another estimate of the image is given by

$$\hat{f}_{2j} = \sum_{i=1}^m A_{ij} \frac{r_i}{L_i}. \quad (2.59)$$

Finally, the adjoint backprojection of (2.54) is approximated as

$$\hat{f}_{Aj} = \sum_{i=1}^m A_{ij} W_{ij} r_i, \quad (2.60)$$

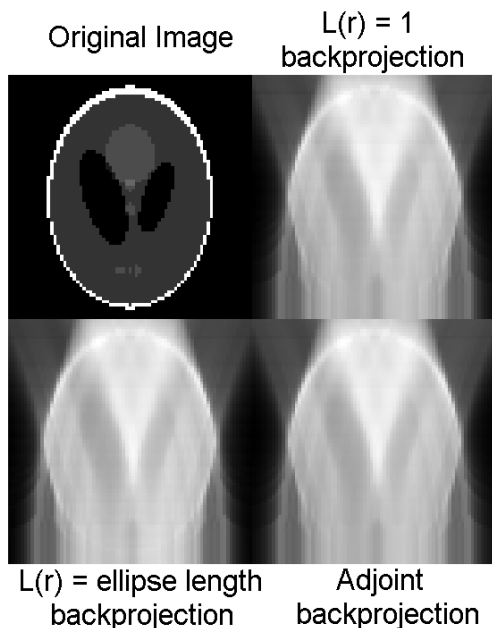


Figure 2.6: Unfiltered backprojections using various weighting strategies for a Shepp-Logan phantom, derived over an aperture of ± 2.5 image widths.

where W_{ij} is the value of W_A appropriate for the i^{th} ellipse and the j^{th} pixel. The value of W_A is approximated using an average of the $x - u$ value within the pixel.

Figure 2.6 compares the image quality of each type of unfiltered backprojection to that of the original image, for a 75×75 pixel Shepp-Logan phantom. The x axis relative to the \mathcal{R}_e transform is coincident with the lower edge of the image. The function $\mathcal{R}_e[f](x, r)$ is sampled in the x dimension at the pixel pitch δ_x over the range of ± 2.5 image widths; the function is sampled in the r dimension, also spaced at δ_x , with limits from $r = \delta/2$ until the ellipses are too large to intersect with any pixel of the test image. The final results of each backprojection are normalized to have the same maximum pixel value.

The results indicate that each backprojection gives a recognizable approxima-

tion of the original image, in which the major features of the Shepp-Logan image are distinguishable. However, the reconstructed images are relatively (and equally) coarse, and do not faithfully reproduce the finer details of the original image. In the case of the normal Radon transform, such behavior is corrected by filtering the backprojection by a function called a Riesz kernel. In principle, a similar filtering is possible for the circular Radon transform; however such filtering is exactly valid only for infinite aperture. It is not clear from the available literature what should be done in the numerical cases which require finite aperture. Perhaps this is the reason for the paucity of numerical results in the literature for that transform.

2.5 The approximate inverse method

The preceding results for the inversion of \mathcal{R}_e add valuable insight and are useful as a connection to the body of literature dealing with the simplest trajectory for down-looking, monostatic SARs and the equivalent problems in integral geometry; furthermore, the backprojections do produce a recognizable, if approximate, representation of the actual image. However, several physical problems limit the direct use of the \mathcal{R}_e trajectory and its formulaic inversion methods, in real applications. These problems include: methods to deal with finite range and sampling in the x and r dimensions; measurement imperfections such as trajectory perturbations, antenna directivity, or ground curvature; the inability to directly apply estimation methods which respect the nature of the measurement noise; and the inability to directly apply the techniques to general multistatic trajectories. Finally, the method cannot truly be said to employ omnidirectional antennas because the inversion can only hope to recover the even part of $f(x, y)$ with respect to the trajectory centerline.

A version of the *approximate inverse* method, developed for tomography prob-

lems [37], is employed as a solution to these concerns. This numerically-based method works simply by writing an equation for each sampling-ellipse measurement in terms of unknown variables which represent the reflectivity function f . The resulting system of equations must be designed such that the unknown variables, representing the image, can be stably estimated. There are few other constraints in using the method: specifically, the equations can be designed to describe any multistatic trajectory, non-flat local geometry, non-uniform antenna sensitivities, anisotropic reflectivity, trajectory perturbations from a nominal path, and many other practical impairments of a real system.

The continuous function $f(x, y)$ is approximated by a discrete spatial convolution:

$$\hat{f}(x, y) = \sum_j \hat{f}_j \delta(x - X_j, y - Y_j) \quad (2.61)$$

where \mathbf{X} and \mathbf{Y} are vectors defining some general grid, and $\hat{\mathbf{f}}$ is a vector whose elements dictate the value of $f(x, y)$ in the vicinity of the point (X_j, Y_j) . The function $\delta(x, y)$ is some two-dimensional approximation to a delta function called a *mollifier*. The mollifier is typically zero at all grid points except at its local origin, and, it defines the interpolation of $\hat{f}(x, y)$ between grid points.

Defining an function approximation error term $\epsilon(x, y) = \hat{f}(x, y) - f(x, y)$, a linear equation in $\hat{\mathbf{f}}$ can be generated for the i^{th} sampling ellipse corresponding to a general measurement $\mathcal{R}[f](x_i, r_i)$. The operator \mathcal{R} is not necessarily \mathcal{R}_e ; the transform coordinates (x_i, r_i) may define the center and semiminor axis of any ellipse, whose orientation and aspect ratio are determined by the trajectories of the general transform. The equation is given by applying \mathcal{R} to both sides of $\hat{f} = f + \epsilon$:

$$\sum_j \hat{f}_j \mathcal{R}[\delta(x - X_j, y - Y_j)] = \mathcal{R}[f] + \mathcal{R}[\epsilon], \quad (2.62)$$

or

$$\sum_j A_{ij} \hat{f}_j = r_i + \eta_i \quad (2.63)$$

where $r_i \triangleq \mathcal{R}[f](x_i, r_i)$ is the measurement, η_i is some hopefully-small disturbance due to the approximation of $f(x, y)$, and the constants $A_{ij} \triangleq \mathcal{R}[\delta(x - X_j, y - Y_j)](x_i, r_i)$ depend on the transform variables (x_i, r_i) , and on the form and offset of the mollifier. If the SAR system parameters and the image resolution are such that η_i may be neglected, then many such equations, representing many sampling ellipses, can be written in matrix form:

$$\mathbf{A}\hat{\mathbf{f}} = \mathbf{r}. \quad (2.64)$$

The matrix \mathbf{A} does not depend on the measurements and therefore may be precomputed for a given set of sampling ellipses.

The approximate inverse is a powerful method which can be simply modified to handle complications in the physical problem which are not quickly amenable to analytical treatment. Multiple synthetic apertures of arbitrary shape merely require rotation extensions of $\mathcal{R}[f]$ to include line integrals on ellipses whose axes do not necessarily lie parallel to the reference axes; otherwise, the formalism derived above is exactly the same.

If the transmit antennas do not realize perfect spherical symmetry, we may simply modify the coefficient matrix \mathbf{A} to represent weighted a weighted average:

$$\mathbf{A}_{ij} = \mathcal{R}[\gamma(x, y)\delta(x - X_i, y - Y_j)](x, r) \quad (2.65)$$

where $\gamma(x, y)$ represents the antenna gain as seen at the point of reflection. (If γ is approximately constant over the support of the mollifier, then it may be taken outside of the integral.) Analytical treatment of this physical situation is the province of the

attenuated Radon transform [40]. Nearly all, if not all, of the idealizations used in the previous analytical section may be relaxed with an equally modest increase in complexity.

However, this power and flexibility comes at a price. A very modest image has perhaps 1000 by 1000 pixels, for which a numerical inversion of a 1,000,000 by 1,000,000 matrix is required. It is easy to imagine requirements for high-resolution images of large areas to involve matrices of dimension in the billions, trillions, or higher. If the mollifier δ is only of local support, then the matrix \mathbf{A} is sparse. Nevertheless, we may expect that the inversion of \mathcal{R} in the general case will be highly demanding from a computing resources point of view. While we hope to examine more-efficient numerical inversion methods in the future of this work, high complexity at the central processor is an expected outcome.

2.5.1 Maximum likelihood image reconstruction

There exist two broad categories of noise classifications in radar systems. Broadband electronic noise present in the low noise amplifiers (LNAs) in the radar receiver is an important and well-studied noise source. Because of system tradeoffs between transmit power and receive SNR, this noise source almost always plays a significant role. The effect is universally approximated as a white Gaussian noise source.

The second category, which captures the signal disturbances due to stochastic phenomenon in the reflection mechanisms of the target or scene, is called *clutter*. The term captures an extremely broad set of phenomena; and because there exist various model approximations for each phenomenon, the literature is replete with distribution functions describing the random processes in a SAR system.

The Rayleigh and Rician densities are used when uncontrolled or random multipath signals exist [41]. Complex Gaussians are used in polarimetric SAR [42]. The

K and β distributions are studied for single and multilook SAR [43, 44, 45]. Most recently, the Γ distribution (a simpler version of the K distribution) has become popular in SAR analysis [46, 47] because of properties we shall discuss below.

Our Radon transform framework essentially assumes that the reflectivity function is always positive; therefore we would expect to use one-sided clutter distributions. In addition, we expect that completely opaque areas with $f(x, y) = 0$ will have very little clutter variance; highly reflective areas, with many reflectors pointing at random angles, should have high variance. Candidate clutter distributions for this application should behave approximately like a Poisson distribution, where the variance and the mean value are linearly related.

Equation 2.64 represents a general discretized relationship between the sampling geometry of $\mathcal{R}[f]$ (represented by the matrix \mathbf{A}), the reflectivity function f , and the measurements \mathbf{R} . The development of a maximum likelihood estimator for the vector \mathbf{f} requires that we model either the elements f_i , or R_i , or both, in a stochastic way. The purpose of this section is to develop the consequences of ML estimation for several types of representative noise phenomenologies. For tractability in the analysis we shall always assume independence in the stochastic processes.

Receiver Noise

In this section, we consider the case in which the measurements R_i are corrupted by i.i.d. Gaussian noise. This model describes the case when clutter noise in the f_i is small enough to treat the f_i variables as deterministic, and electronic noise at the receiver is a dominant contributor to the total disturbance. Then, we have

$$R_j = \sum_i A_{ij} f_i + n_j = \bar{R}_j + n_j \quad (2.66)$$

such that $R_j \sim \mathcal{N}(\sum_i A_{ij} f_i, \sigma^2)$. The index i varies over all cells in the problem. Then

$$p(\mathbf{R}|\mathbf{f}) = \prod_j \mathcal{N}_{R_j}(\sum_i A_{ij} f_i, \sigma^2) \quad (2.67)$$

where j indexes all the measurements R_j . The ML condition for each of the cells, indexed by k , leads directly to a corresponding equation which must be satisfied:

$$\frac{d \ln p(\mathbf{R}|\mathbf{f})}{df_k} = 0 \Rightarrow \sum_{j=1}^N A_{kj} \sum_{i=1}^N A_{ij} \hat{f}_i = \sum_{j=1}^M A_{kj} R_j, \quad 1 \leq k \leq N \quad (2.68)$$

representing an N by N square system of linear equations which may be recognized as the normal regression equations in matrix form:

$$\mathbf{A}^T \mathbf{A} \hat{\mathbf{f}} = \mathbf{A}^T \mathbf{R} \quad (2.69)$$

which has a solution

$$\hat{\mathbf{f}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{R} \quad (2.70)$$

if \mathbf{A} is well-conditioned.

Receiver noise with path loss

As first noted by Norton [27], the measurements represented by the forward Radon transform are actually subject to a path loss which is not the same at all points. At any given point in the measurement space, all signals collected at the corresponding time instant must have traveled an equal distance and therefore may be said to undergo the same path loss. Assuming this path loss function is a known function of the distance r then we may choose to re-normalize measurements by dividing out

the loss function such that again

$$R_j = \sum_i A_{ij} f_i + n_j = \bar{R}_j + n_j, \quad (2.71)$$

however the variance of the added noise is now a function of the total travel distance K . We capture this effect by defining the variance $\sigma^2(K_j)$, an increasing function of K . In the simulations (below) we shall use the form $\sigma^2(K) = \sigma_0^2(K/K_{min})^\gamma$. The measurement is distributed as $R_j \sim N(\sum_i A_{ij} f_i, \sigma^2(K_j))$. Similar calculations again lead to a linear system of N equations in the N unknowns:

$$\sum_{j=1}^N \frac{A_{kj}}{\sigma^2(K_j)} \sum_{i=1}^N A_{ij} \hat{f}_i = \sum_{j=1}^M \frac{A_{kj}}{\sigma^2(K_j)} R_j, \quad 1 \leq k \leq N \quad (2.72)$$

which is a weighted least-squares formulation.

Rayleigh measurement noise

We now postulate that the measurements are Rayleigh distributed with mean given by the non-random variable $\bar{R}_j = \sum_i A_{ij} f_i$. The Rayleigh distribution admits positive random variables whose means are proportional to their standard deviations. The independent measurements give the stochastic model

$$p(\mathbf{R}|\bar{\mathbf{R}}) = \prod_{j=1}^M \frac{\pi R_j e^{-\frac{\pi R_j^2}{4\bar{R}_j}}}{2\bar{R}_j^2} \quad (2.73)$$

Taking the logarithm and then the derivative with respect to each of the f_k , yields the system of N nonlinear equations

$$\sum_{j=1}^M \frac{A_{kj}}{\sum_i^N A_{ij} \hat{f}_i} = \frac{\pi}{8} \sum_{j=1}^M \frac{A_{kj} R_j^2}{[\sum_i^N A_{ij} \hat{f}_i]^2}, \quad 1 \leq k \leq N \quad (2.74)$$

The sufficient statistics are dependent on the square of each measurement, R_j^2 . Numerical instability is likely with this ML scheme for measurements which are near zero under this model, because the equations blow up under these conditions. The meaning of a zero measurement is that all cells which are touched by the line integral *must* be zero. Such cells and measurements can be deleted from the problem. The remaining coefficients in the equations will be finite.

Gamma-distributed clutter noise

The final stochastic model considered here is one in which the f_i themselves are treated as random variables from measurement to measurement. The problem strategy in this case is to estimate the mean, \hat{f}_i from the measurements R_j .

In general it is difficult to generate the density function $p(\mathbf{R}|\bar{\mathbf{f}})$ because R_j is the sum of many random variables f_i . However, under a certain restriction, the gamma distribution allows very simple construction of $p(\mathbf{R}|\bar{\mathbf{f}})$ [46].

The gamma distribution takes two parameters, α and β , giving a density function

$$p(x; \alpha, \beta) = \frac{x^{\alpha-1} \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)}, \quad x \geq 0. \quad (2.75)$$

The mean of this distribution is $\mu = \alpha/\beta$, and the variance is $\sigma^2 = \alpha/\beta^2$. The mean is proportional to the variance, like its discrete counterpart, the Poisson distribution.

If β is fixed, and if independent random variables are distributed $X_i \sim \Gamma(\alpha_i, \beta)$, then $\sum X_i \sim \Gamma(\sum \alpha_i, \beta)$. This property is most easily seen by noting that the characteristic function is $\theta(\tau) = (1 - i\tau/\beta)^{-\alpha}$. Therefore, if $f_i \sim \Gamma(\bar{f}_i, \beta)$, then

$$p(\mathbf{R}|\bar{\mathbf{f}}) = \prod_{j=1}^M \frac{R_j^{\beta \sum_i A_{ij} \bar{f}_i} \beta^{\beta \sum_i A_{ij} \bar{f}_i} e^{-\beta R_j}}{\Gamma(\beta \sum_i A_{ij} \bar{f}_i)} \quad (2.76)$$

Taking the log and the derivative with respect to \bar{f}_k as usual, we can generate a set of N nonlinear equations which must be satisfied with the ML estimate $\hat{\mathbf{f}}$:

$$\sum_{j=1}^M A_{kj} \Psi\left(\beta \sum_i A_{ij} \hat{f}_i\right) = \sum_{j=1}^M A_{kj} (\ln R_j + \ln \beta), \quad 1 \leq k \leq N \quad (2.77)$$

where $\Psi(x)$ is the digamma function. In this case, the sufficient statistics are based on $\ln \mathbf{R}$. We note that if $R_j = 0$, then, as with the Rayleigh case, all cells touching the ellipse of integration are identically zero.

2.6 Numerical approximate inverse results

This section contains simulations of the least-squares solution (2.70) for a Shepp-Logan test image. The vectors \mathbf{X} and \mathbf{Y} are chosen to form an N by N rectangular grid spaced at pixel dimension δ_x . A very simple mollifier is chosen such that the offset mollifiers exactly tile the rectangular image:

$$\delta(x, y) = \text{rect}\left(\frac{x}{\delta_x}\right) \text{rect}\left(\frac{y}{\delta_x}\right) \quad (2.78)$$

where

$$\text{rect}(x) = \begin{cases} 1 & \text{if } |x| < \frac{1}{2}, \\ 0 & \text{if } |x| > \frac{1}{2}. \end{cases} \quad (2.79)$$

The resulting $\hat{f}(x, y)$ approximation resembles a checkerboard. The integrals defining \mathbf{A}_{ij} are just the lengths of the elliptical arc through the square pixel surrounding the point (X_i, Y_i) . Calculation of these arc lengths require, in general, multiple evaluations of an incomplete elliptic integral of the second kind.

Figure 2.7 shows the approximate inverse results for \mathcal{R}_e under two conditions of

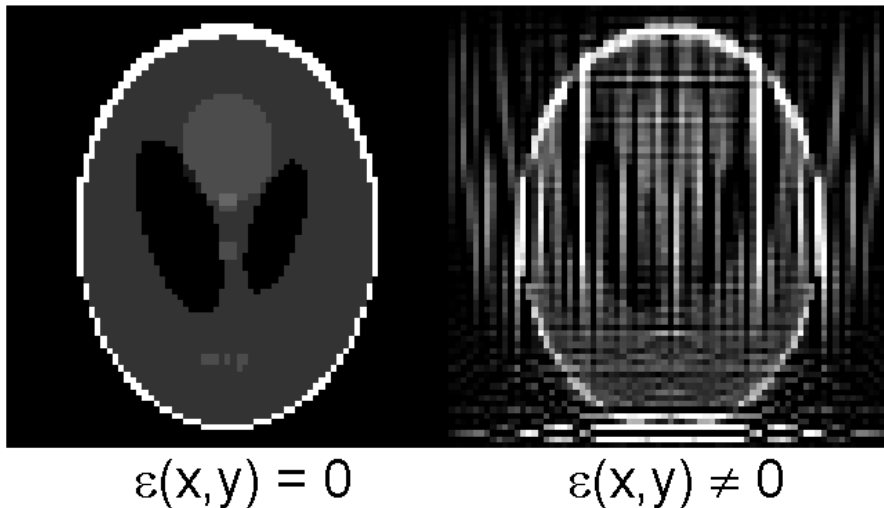


Figure 2.7: Comparison of 75×75 approximate inverse reconstructions of the \mathcal{R}_e trajectory without (left) and with (right) function-approximation error.

function-approximation error. In the left-hand image, there is no function-approximation error, that is, the measured function $f(x, y)$ actually *is* a Shepp-Logan checkerboard function expressible exactly in the form $\hat{f}(x, y)$. In the right-hand image, another extreme is represented in which the measurement vector \mathbf{r} has nothing to do with the image discretization, but measures the exact line integrals of the exact shapes comprising a Shepp-Logan image. In this case, the measurement is of infinite bandwidth, giving rise to the greatest ambiguities when a sampling ellipse is parallel to an edge of $f(x, y)$. The \mathbf{r} vectors are the only difference between the two cases. The common m by n matrix \mathbf{A} represents $m = 33,600$ measurement ellipses for an $n = 75 \times 75$ pixel image over an aperture of ± 2.5 image widths. All other sampling details are the same as described above for the backprojection simulations.

The comparison illustrates that \mathbf{A} is well-conditioned enough to give an essentially noiseless reconstruction when the disturbances η_i are zero. On the other hand, the

poor image quality using the infinite-bandwidth measurement shows that either the η_i are significant, or the matrix \mathbf{A} is ill-conditioned. In fact, both are factors, but by far the highest impact comes from the condition number of \mathbf{A} : a large fraction of the singular values of \mathbf{A} are very near to zero, such that perturbations in \mathbf{r} are magnified in the inversion. The reasons for this behavior are inherent in the nature of the \mathcal{R}_e trajectory, as shown in another section of the present work.

Finally, the reconstruction from a multistatic trajectory of approximately the same number of sampling ellipses is shown. The multistatic trajectory is a *random* trajectory with bistatic pairs whose locations are randomly drawn from a uniform distribution which is $h = 1$ to $h = 2$ image widths above the image area, giving ellipses whose orientation is random with respect to the reference axes. Each pair contributes one transmit burst's worth of sampling ellipses. The sampling pitch and range details are equivalent to those used previously. This trajectory removes most limitations of the \mathcal{R}_e trajectory with respect to the omnidirectional-antenna condition; in particular, the trajectory eliminates the limitation due to the null space of \mathcal{R}_e . The remaining limitation (that of sampling ellipses exceeding the bounds of the image) is the subject of a concurrent work. Figure 2.8 compares the performance of this random multistatic trajectory to that of \mathcal{R}_e .

While the multistatic reconstruction has noticeable imperfections, which are due to finite function approximation error effects, it is a clear improvement over that of the simple bistatic trajectory. In this case, the distribution of singular values of \mathbf{A} shows no values very near to zero. Therefore, this inversion is much more robust against disturbances due to function approximation error, as well as superior for i.i.d. Gaussian measurement noise.

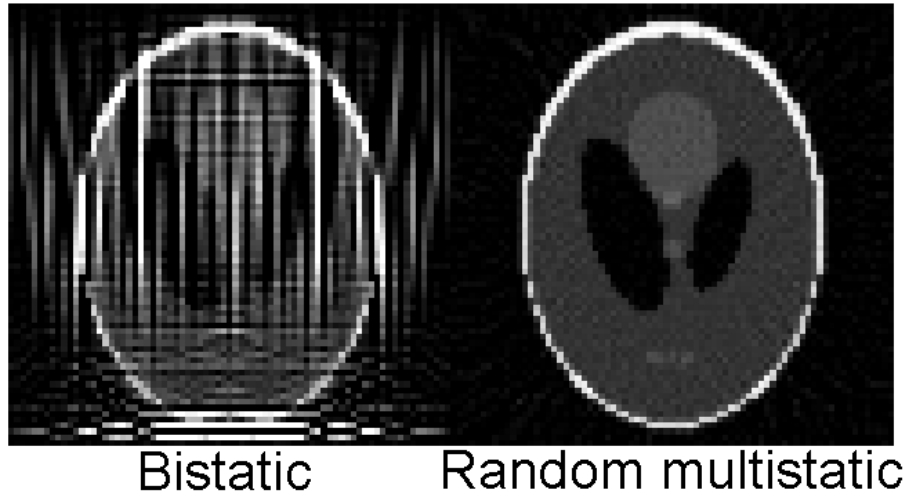


Figure 2.8: Comparison of 75×75 approximate inverse reconstructions of the \mathcal{R}_e trajectory (left) and a random multistatic trajectory (right). Both measurements exhibit finite function-approximation error.

2.7 Conclusions

The integral transform framework of down-looking monostatic radars, typified by the circular Radon transform and its simple straight-line trajectory, has been generalized to treat the general case of downlooking multistatic radars. Such radars can be thought of as measuring the line integrals around an expanding set of ellipses, which are nontrivial functions of the loci of the transmitters and receivers. The relationships between the ellipse parameters and the radar system geometries were developed for general multistatic trajectories.

The simplest such multistatic trajectory was described by an elliptical integral transform. A general frequency-domain backprojection method was developed for approximate inversions of the transform. These backprojections involve reasonably-complicated Fourier and Hankel transformations of weighted versions of the forward

transforms. The adjoint operator of this transform was found, and its inversion performance was compared to simpler, more intuitive backprojection methods.

Explicit inversions of the elliptical transform, such as exist for the circular transform, were not found in this work due to intractabilities arising from the variation of the aspect ratio of the ellipses with their minor axis. Future researchers might begin with the adjoint operation developed here, and attempt to develop a filtered backprojection algorithm based on the adjoint. The successful demonstration of the existence of such an inversion may not be trivial.

The approximate inverse method exhibited far superior inversion performance, and adaptability to a number of non-ideal measurement phenomena. The method reveals that the simplest bistatic measurement scheme is inherently ill-conditioned, in addition to the scheme's other limitations. Changing the sampling strategy to a random multistatic scheme showed relatively excellent inversion performance, in addition to relieving many of the simple bistatic scheme's limitations.

Chapter 3

Performance synthesis for UAV trajectories in multistatic SAR

Trajectory design in Unmanned Aerial Vehicle (UAV) systems has taken a number of interesting forms in the literature. The wide range of results and techniques arise from the associated variety of design goals. A sparse sampling of the field includes the design of multiple-UAV trajectories for the purposes of collision avoidance [48], minimum radar exposure [49], and terrain avoidance [50]. UAV trajectories may also be optimized to achieve a specific signal processing requirement. For example, [51] provides performance bounds the problem of optimal ground-moving-target-indicator (GMTI) performance for a group of Doppler-radar-carrying UAVs, and finds an optimal UAV geometric configuration which achieves these bounds.

In this work, we address the question of UAV trajectory design for the purpose of robust image performance in a multistatic SAR system. A general configuration of strong recent interest is that of multiple UAVs operating more or less in coordination over or nearby a scene of interest. The constraint of hardware simplicity can drive the design of such vehicles to use simple, relatively-omnidirectional an-

tennas instead of more complex beamformers. It is therefore of interest to study optimal image-reconstruction techniques under these constraints. We consider the image-reconstruction performance of the general class of multistatic trajectories, and examine examples in the random-sampling trajectory, and deterministic trajectory, subclasses.

The contribution of this work falls in several categories. First, a novel noise figure of merit, based on a tomographic model of the SAR measurement process, provides a guiding metric for trajectory optimization in the presence of measurement noise. Results of modern matrix analysis are applied to generate an absolute performance bound, which we show to be achievable, but only by an uninteresting type of trajectory. Connections of this trajectory-design problem with the Welch bound of communications theory are shown, giving useful design guidelines from that discipline. Translation of these mathematical results to physical parameters of the SAR system and measurement provides an intuitive physical interpretation of trajectories which produce good performance potential. Several examples of multistatic trajectories are simulated and compared to these bounds. Finally, we develop asymptotic upper bounds from random matrix theory which predict the performance of large SAR problems quite well; furthermore, because this bound asymptotically achieves our lower bound, this work suggests that random trajectories will perform equal to or better than any deterministically-designed trajectory, for large enough problems.

Image reconstruction using random sampling is also explored in the field of generalized or irregular sampling [52, 40], and in *compressed sensing* (CS) [53, 54]. These methods reconstruct, from some sort of irregular sampling or projections, an image which is sparse in some *known* basis. For example, in very recent and important work [55], Pan successfully shows image reconstruction from randomly-placed sensors which directly and noiselessly sample local image pixels, wherein the image

processor interpolates the unmeasured pixels by an advanced iterative reconstruction algorithm. The present work employs a similar uniformly-distributed strategy for randomly placing the UAV “sensors”. The detail of the sampling method in multistatic SAR, which is a generalized projection and not a direct sampling, provides a measure of contrast to the present work. However, again the fundamental contrast is the treatment of measurement noise. Compressed sensing researchers have shown that good reconstruction should be expected using noiseless random projections of any image. The present work shows that, for feasible SAR measurements, randomly-placed sensors provide not only good reconstruction using noiseless measurements, but also indicates that such reconstructions can generate the best possible performance of standard maximum likelihood inversions, *in the presence of noise*, of any multistatic UAV trajectory, under certain technical conditions.

In the bulk of the present work, no fundamental assumption is made about the image details. In particular, unlike CS, the image is not assumed to be sparse in some basis. In the final section, however, we address the question of whether the basic advantages of CS may remain applicable under the constraints on the sampling projections created by the SAR measurement process.

The chapter is organized as follows. In section 3.1, a SAR measurement framework is reviewed and a noise figure-of-merit is derived for a maximum-likelihood estimate of the image based on such measurements. Sections 3.2, 3.3, and 3.4 develop absolute and pragmatic bounds on this noise figure and relate these bounds to physical properties of the SAR measurements. The resulting guidelines for the design of the multistatic SAR trajectories are summarized in section 3.5. We give several examples of multistatic SAR trajectories and compare their noise performance with expectations in section 3.6. The power of random matrix theory is brought to bear in order to estimate the performance of random-location trajectories in section 3.7. The

problem of edge effects and practical sampling constraints is discussed and resolved by the use of variable-resolution pixels in section 3.8. In section 3.9, we demonstrate that the concepts of random-projection compressed sensing remain in force in the presence of physical SAR sampling constraints. Conclusions and final remarks are found in section 3.10.

3.1 A noise figure performance metric

Due to the complexity of the measurement and fluctuation phenomenon in a real SAR system, the characterization of its image-reconstruction performance by a single number is a necessarily approximate endeavor. The figure of merit used here retains the virtues of simplicity and amenability to the power of matrix analysis, while capturing (albeit approximately) some of the complexity of a SAR imaging system.

The SAR measurement framework used here is a version of the *approximate inverse*, developed for downlooking multistatic SAR in [56]. These results can be briefly summarized as follows.

In one transmission burst, an omnidirectional, bistatic receiver/transmitter pair, over a flat plane of reflectivity function f , generates a series of measurements which are the line integrals of f over an expanding set of ellipses. The geometry of these ellipses depend on the coordinates of the UAVs above the plane.

A typical example of these sampling ellipses (Figure 3.1) shows several measurement ellipses generated during a single transmit burst between a bistatic pair. Each ellipse is the locus of points, in the reflectivity plane, for which photons leaving the transmitter and gathered at the receiver have traveled an equal distance. The physical measurement (that is, the signal strength at the receiver at a particular instant corresponding to a given ellipse) is modeled as the line integral of the reflectivity

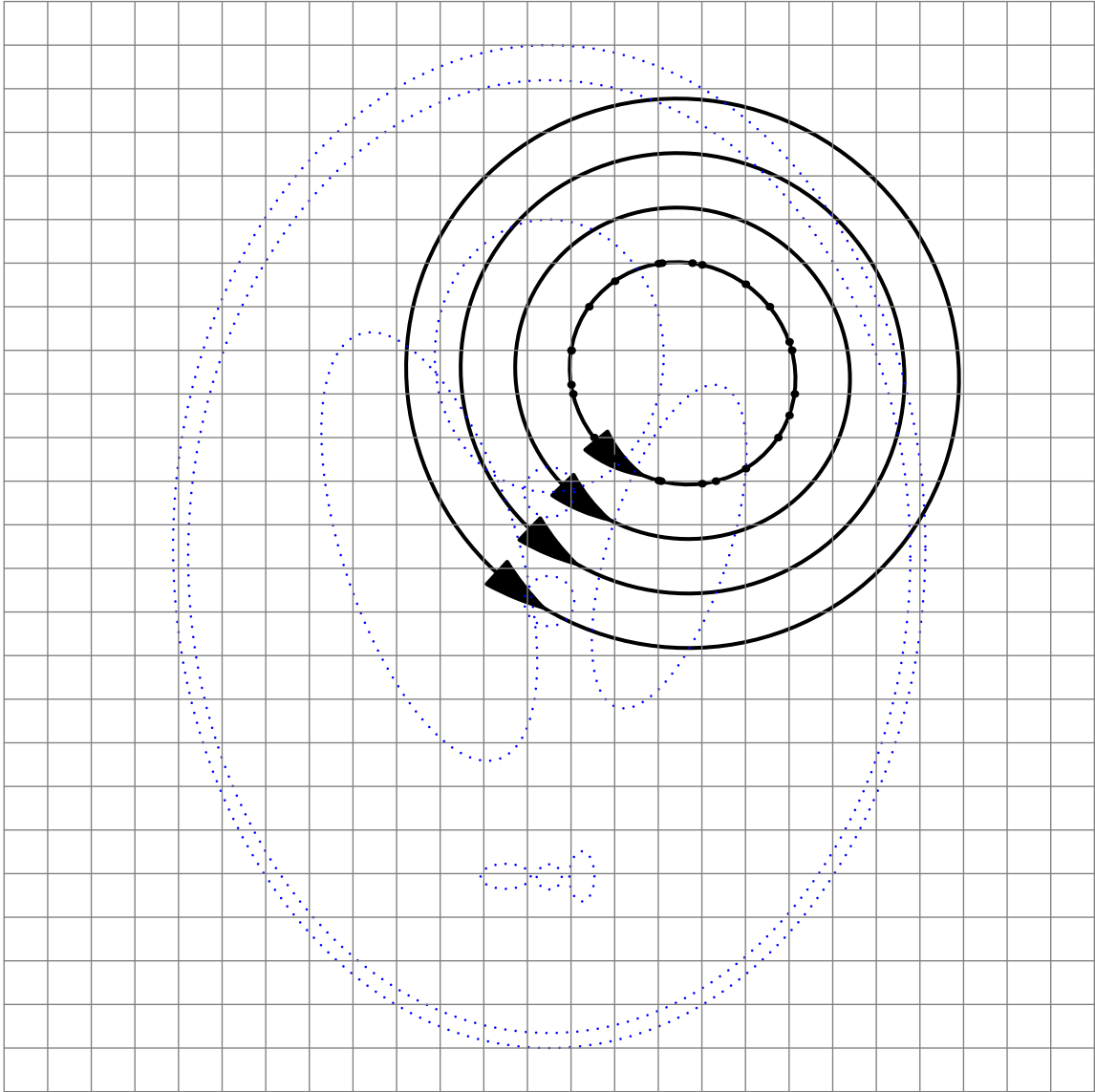


Figure 3.1: A representative example of four ellipses of integration, each representing a row of the matrix \mathbf{A} . The light dotted curves represent contours of some reflectivity function $f(x, y)$.

function $f(x, y)$ around the ellipse. That measurement is an element of the measurement vector \mathbf{r}_e . Using the simplest possible discretization of $f(x, y)$, the function $f(x, y)$ is approximated by a checkerboard function $\hat{f}(x, y)$, which is constant within each grid cell, and whose value within that cell is specified by an element of the vector $\hat{\mathbf{f}}$. In such a case, an approximation of the line integral is the sum of $\hat{f}_j A_{ij}$, where A_{ij} is the segment length of the i^{th} sampling ellipse passing through the j^{th} grid cell corresponding to \hat{f}_j . By setting this approximation of the line integral to be equal to the corresponding radar measurement in \mathbf{r}_e , for many measurement ellipses arising from (possibly) many such bistatic pairs, the SAR measurement process is simply described as

$$\mathbf{A} \hat{\mathbf{f}} = \mathbf{r}_e, \quad (3.1)$$

where \mathbf{A} comprises the known A_{ij} values and is an m -measurement by n -pixel measurement matrix, $\hat{\mathbf{f}}$ is an n -pixel vector of unknown pixels comprising the image, and \mathbf{r}_e is the known measurement vector of length m . Each component of \mathbf{r}_e measurement vector is therefore a projection of \mathbf{f} onto the rows of \mathbf{A} .

Clear tradeoffs exist between the dimension of the cell δ_x , the fidelity of the approximating function $\hat{f}(x, y)$, and the complexity of the inversion process. We note that while it is possible in an infinite-bandwidth measurement system to measure the independent line integrals of infinitely many ellipses in a single transmit burst, in practice a real finite-bandwidth system provides finite information which is representable by a sampling of the processed waveform. We shall assume that the measurement hardware has sufficient bandwidth to support the pixel resolution δ_x ; furthermore, we shall increment the axes of successive sampling ellipses by the amount δ_x .

If measurement noise in each element of \mathbf{r}_e can be thought of as i.i.d. Gaussian

disturbances, then a maximum likelihood estimate of the values $\hat{\mathbf{f}}$ is the standard least-squares solution

$$\hat{\mathbf{f}}_{ML} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{r}_e \quad (3.2)$$

$$= \mathbf{A}^+ \mathbf{r}_e \quad (3.3)$$

$$= \mathbf{V} \mathbf{\Sigma}^+ \mathbf{U}^T \mathbf{r}_e, \quad (3.4)$$

where \mathbf{A}^+ is the pseudoinverse of \mathbf{A} , \mathbf{U} and \mathbf{V} are orthogonal matrices, and $\mathbf{\Sigma}^+$ is a diagonal matrix whose elements are given by the reciprocals of the singular values on the diagonal of the matrix $\mathbf{\Sigma}$ in the singular value decomposition $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$. If the estimate $\hat{\mathbf{f}}_{ML}$ is to remain unbiased, then all components of the output space spanned by the columns of \mathbf{V} must be retained.

If the measurement vector is decomposed into its signal and zero-mean noise component, $\mathbf{r}_e = \bar{\mathbf{r}}_e + \mathbf{n}$, then the noisy component of the reconstructed image is

$$\mathbf{f}_N = \mathbf{V} \mathbf{\Sigma}^+ \mathbf{U}^T \mathbf{n}. \quad (3.5)$$

An appropriate statistic for the noise image vector \mathbf{f}_N is the average variance of its pixels. This quantity is derived by tracing the projections of \mathbf{n} through the multiplications in equation 3.5. If the components of \mathbf{n} are zero-mean, i.i.d. Gaussian random variables of variance σ_I^2 then the initial multiplication by the orthonormal matrix \mathbf{U}^T produces another random vector whose statistics are indistinguishable from those of \mathbf{n} . Multiplication of the result by the diagonal matrix $\mathbf{\Sigma}^+$ produces n non-zero independent Gaussian variables of generally-unequal variance σ_I^2/s_i^2 , for $i \leq n$, where s_i is the i^{th} singular value of \mathbf{A} . The final multiplication by \mathbf{V} produces

random variables in the i^{th} pixel of variance

$$\sigma_i^2 = \sigma_I^2 \sum_{j=1}^n V_{ij}^2 \frac{1}{s_j^2}, \quad (3.6)$$

which is generally different for each pixel. A more useful statistic, which does not depend on the details of \mathbf{V} , is obtained by a simple average of these variances:

$$\frac{1}{n} \sum_{i=1}^n \sigma_i^2 = \frac{1}{n} \sigma_I^2 \sum_{i=1}^n \sum_{j=1}^n V_{ij}^2 \frac{1}{s_j^2} \quad (3.7)$$

$$= \frac{1}{n} \sigma_I^2 \sum_{j=1}^n \frac{1}{s_j^2} \sum_{i=1}^n V_{ij}^2 \quad (3.8)$$

$$= \frac{1}{n} \sigma_I^2 \sum_{j=1}^n \frac{1}{s_j^2} \triangleq \sigma_O^2, \quad (3.9)$$

where the last step follows from the normality of \mathbf{V} . Therefore, we may define a noise figure, which is the ratio of the output image noise variance to the input measurement variance:

$$G = \frac{\sigma_O^2}{\sigma_I^2} = \frac{1}{n} \sum_{i=1}^n \frac{1}{s_i^2} = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i}. \quad (3.10)$$

Here, λ_i is an eigenvalue of $\mathbf{A}^T \mathbf{A}$, which equals the square of an associated singular value, s_i^2 , of \mathbf{A} . The number G is a function only of the geometry of the measurement strategy, as captured in the projection matrix \mathbf{A} , and does not depend on the image detail. By convention, the eigenvalues are ordered such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_n \geq 0$.

The number G is identical to *least-squares efficiency* [57] and is known in many other guises, e.g., $G = \frac{1}{n} \text{Tr}((\mathbf{A}^T \mathbf{A})^{-1})$. A variety of applications use a similar figure of merit [58, 59, 60]. In the following, we collate matrix analysis results from such work, in addition to the theory of positive matrices, to illuminate the multistatic

SAR inversion problem.

3.2 Absolute performance bounds

A version of the well-known Kantorovich inequality (see, for example, [61]) states that for an ordered sequence λ_i and any weights α_i such that $\sum \alpha_i = 1$,

$$1 \leq \sum \alpha_i \lambda_i \sum \frac{\alpha_i}{\lambda_i} < \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}. \quad (3.11)$$

Furthermore, because the trace of any square matrix is the same as the sum of its eigenvalues, the average eigenvalue of $\mathbf{A}^T \mathbf{A}$ in terms of statistics of \mathbf{A} is

$$\bar{\lambda} = \frac{1}{n} \text{Tr}(\mathbf{A}^T \mathbf{A}) = \frac{1}{n} \|\mathbf{A}\|_F^2, \quad (3.12)$$

where the Frobenius norm is used. Because the unweighted sum of $1/\lambda_i$ is of interest, we set $\alpha_i = 1/n$ to achieve the result

$$\frac{1}{\bar{\lambda}} \leq \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i} \leq \frac{(\lambda_1 + \lambda_n)^2}{4\bar{\lambda}\lambda_1\lambda_n} = \frac{1}{4\bar{\lambda}}(\kappa + 2 + \kappa^{-1}), \quad (3.13)$$

where $\kappa = \lambda_1/\lambda_n$ is the 2-norm condition number of $\mathbf{A}^T \mathbf{A}$. Equation 3.13 gives upper and lower bounds for the noise figure G . All inequalities become equalities when all $\lambda_i = \bar{\lambda}$, which is the only way to achieve the lower bound. The upper bound is matched when the upper half of the eigenvalues are equal to λ_1 and the remaining eigenvalues are equal to λ_n . For good condition numbers κ , equation 3.13 gives a tight estimate of G without reference to the distribution of the eigenvalues between λ_1 and λ_n ; for example, $\kappa = 2$ gives upper and lower bounds which are different by a factor of 1.125. However, for ill-conditioned matrices the upper bound is typically

quite pessimistic because it assumes a worst-case bimodal distribution of λ_i , knowing no other statistic of the distribution (such as, its variance).

The minimum possible value of G for any matrix \mathbf{A} is therefore

$$G_{MIN}(\mathbf{A}) = (\bar{\lambda})^{-1} = \frac{n}{\|\mathbf{A}\|_F^2}, \quad (3.14)$$

which is achievable only when the variance of the eigenvalues is zero.

3.2.1 Conditions for meeting the lower bound

The variance of the eigenvalues, σ_λ^2 , can be determined by measuring the trace of $(\mathbf{A}^T \mathbf{A})^2$, which gives the sum of the squares of λ_i :

$$\overline{\lambda^2} = \frac{1}{n} \sum_{i=1}^n \lambda_i^2 = \frac{1}{n} \text{Tr}((\mathbf{A}^T \mathbf{A})^2) = \frac{1}{n} \|\mathbf{A}^T \mathbf{A}\|_F^2, \quad (3.15)$$

and therefore the eigenvalue variance is given by

$$\sigma_\lambda^2 = \overline{\lambda^2} - \bar{\lambda}^2 = \frac{1}{n} \|\mathbf{A}^T \mathbf{A}\|_F^2 - \frac{1}{n^2} \|\mathbf{A}\|_F^4. \quad (3.16)$$

Clearly, σ_λ^2 is minimized when $\overline{\lambda^2}$ is minimized. If $\overline{\lambda^2}$ is written alternatively as

$$\overline{\lambda^2} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (c^{ij})^2, \quad (3.17)$$

where c^{ij} is the correlation of the i^{th} and j^{th} columns of \mathbf{A} , then it is clear that Welch's inequality [62] is applicable to this problem. Welch's explicit result is usually given assuming that the rows of \mathbf{A} have equal energy. However, his result can be modified to give a lower bound for $\overline{\lambda^2}$ without constraining the rows of \mathbf{A} to represent equal ellipse lengths, as seen in Massey's elegant proof of the Welch bound [63]. The result

is

$$\text{Tr}((\mathbf{A}^T \mathbf{A})^2) = \sum_{i=1}^n \sum_{j=1}^n (c^{ij})^2 \geq \frac{1}{n} \|\mathbf{A}\|_F^4, \quad (3.18)$$

with equality if and only if the columns of \mathbf{A} are mutually orthogonal *and furthermore* have equal energy. Comparison of equations 3.16 and 3.18 show that σ_λ^2 is zero if and only if \mathbf{A} meets the modified Welch bound in equation 3.18 with equality. Such matrices are the only matrices which will match the G_{MIN} bound.

There exists a large body of work on sequence sets (that is, candidate rows of \mathbf{A}) which meet or approximate the Welch bound. However, in this application, \mathbf{A} is constrained to be a positive matrix (that is, $A_{ij} \geq 0$) because each element represents the length of an ellipse segment through the corresponding grid element. This constraint is severe; the only positive matrices strictly meeting the Welch bound are square matrices ($m = n$) which are proportional to a permutation matrix. Nevertheless, the Welch equality conditions give sensible insight into the design of the \mathbf{A} matrix. We shall return to this point in section 3.5.

3.3 Pragmatic performance bounds

In most practical cases, the eigenvalue variance σ_λ^2 will be large enough for the image reconstruction performance to stray significantly from the G_{MIN} performance bound. The root cause for this discrepancy may be an inability to match the Welch conditions, or, the designer elects to use easily-included extra sample ellipses (because they are available in a transmit burst), or, simply that reconstruction must proceed only with available data. In these cases, it is useful to have a lower bound based on simply-computed statistics of \mathbf{A} , which respects a finite σ_λ^2 .

3.3.1 Estimates of the dominant eigenvalue

The theory of positive matrices, so constraining in matching the Welch bound, is now very useful in estimating the dominant eigenvalue λ_1 of $\mathbf{A}^T\mathbf{A}$, which is positive because \mathbf{A} is positive. The Perron-Frobenius result [64] states that for general positive matrices, λ_1 is distinct and

$$\min r_i \leq \lambda_1 \leq \max r_i, \quad (3.19)$$

where r_i is the sum of the i^{th} row of $\mathbf{A}^T\mathbf{A}$. The result suggests that a reasonable estimate for λ_1 is the average value of r_i , which clearly lies between the given limits. Indeed Merikoski has shown [65] that the average row sum is a lower bound for positive and symmetric matrices (such as $\mathbf{A}^T\mathbf{A}$). Merikoski also notes that the average row sum appears to be an excellent estimate of the dominant eigenvalue for correlation matrices. Our numerical experience supports this observation. An estimate of the dominant eigenvalue of $\mathbf{A}^T\mathbf{A}$ is therefore:

$$\hat{\lambda}_1 = \frac{1}{n} \sum_{i=1}^n r_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (\mathbf{A}^T\mathbf{A})_{ij} = \frac{1}{n} \|\mathbf{A}^T\mathbf{A}\|_1, \quad (3.20)$$

where the entrywise one-norm $\|\cdot\|_1$ is valid because all entries are non-negative.

3.3.2 An improved lower bound for finite σ_λ^2

In [58], Bai and Golub develop upper and lower bounds for the trace of the inverse of a correlation matrix when the second order statistics of λ_i are known. The bounds are based on $\bar{\lambda}$, $\overline{\lambda^2}$ (both measured directly from the matrix) and estimates for the extreme eigenvalues λ_1 and λ_n . We focus on the lower bound result, which not a

function of the smallest eigenvalue λ_n , and simplifies in this case to

$$G_P(\mathbf{A}) = \frac{1}{\hat{\lambda}_1} \left\{ \frac{(\bar{\lambda} - \hat{\lambda}_1)^2}{\hat{\lambda}_1 \bar{\lambda} - \bar{\lambda}^2} + 1 \right\} \quad (3.21)$$

$$\approx \frac{\left(\frac{\bar{\lambda}}{\hat{\lambda}_1} - 1\right)^2}{\bar{\lambda} - \frac{\bar{\lambda}^2}{\hat{\lambda}_1}} \quad (3.22)$$

$$\approx \frac{1}{\bar{\lambda} - \frac{\bar{\lambda}^2}{\hat{\lambda}_1}}, \quad (3.23)$$

where in the first approximation, the unity term can be dropped because the added value is less than the value G_{MIN} , which in turn is assumed to be much smaller than G_P . The second approximation is valid whenever the dominant eigenvalue is much larger than the mean eigenvalue, which is typically the case in practice. The lower bound G_P becomes G_{MIN} when $\hat{\lambda}_1 = \bar{\lambda}$. The G_P bound provides an efficiency comparison to the measured value of G for candidate \mathbf{A} designs when the Welch bound is not well approximated.

Having generated performance bounds in terms of statistics of the \mathbf{A} matrix, we can now relate these bounds to physical parameters of the SAR measurement.

3.4 Physical interpretation of performance bounds

Under the inversion conditions assumed in this work (that is, the checkerboard approximation of the reflectivity function f), each of the preceding bounds can be related to physical parameters of the SAR process.

3.4.1 Preliminaries

The dominant eigenvalue estimate is simply related to the lengths of the sampling ellipses by

$$\hat{\lambda}_1 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (\mathbf{A}^T \mathbf{A})_{ij} \quad (3.24)$$

$$= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^m A_{ki} A_{kj} \quad (3.25)$$

$$= \frac{1}{n} \sum_{k=1}^m \left(\sum_{i=1}^n A_{ki} \right) \left(\sum_{j=1}^n A_{kj} \right) \quad (3.26)$$

$$= \frac{1}{n} \sum_{k=1}^m E_k^2 \triangleq \frac{\overline{E^2}}{\beta}, \quad (3.27)$$

where E_k is the sum of the k^{th} row of \mathbf{A} (which is the length of the k^{th} measurement ellipse), $\overline{E^2}$ is the average square of the ellipse lengths, and $\beta = n/m$ is the aspect ratio of the matrix \mathbf{A} . Evidently, the dominant eigenvalue depends solely on the lengths of the measurement ellipses and not on the pixel dimension δ_x .

The exact relationship between the ellipse parameters and the mean eigenvalue $\bar{\lambda}$ is complicated by the interplay of the cell boundaries and the sampling phase of the ellipses. However, the general relationship can be understood by the following argument. First, consider an ellipse of length E_i which is described by a row of \mathbf{A} . The ellipse is divided into a number of segments; the number of segments is approximately proportional to E_i/δ_x . The average square of each segment is another number of scale δ_x^2 . Therefore, the sum of the squares of the i^{th} row of \mathbf{A} might be expected to be

$$\sum_{j=1}^n (A_{ij})^2 \approx \alpha_1 \delta_x^2 \frac{E_i}{\delta_x} = \alpha_1 \delta_x E_i. \quad (3.28)$$

Figure 3.2 shows relationship between E_i and the sum of the squares of the cor-

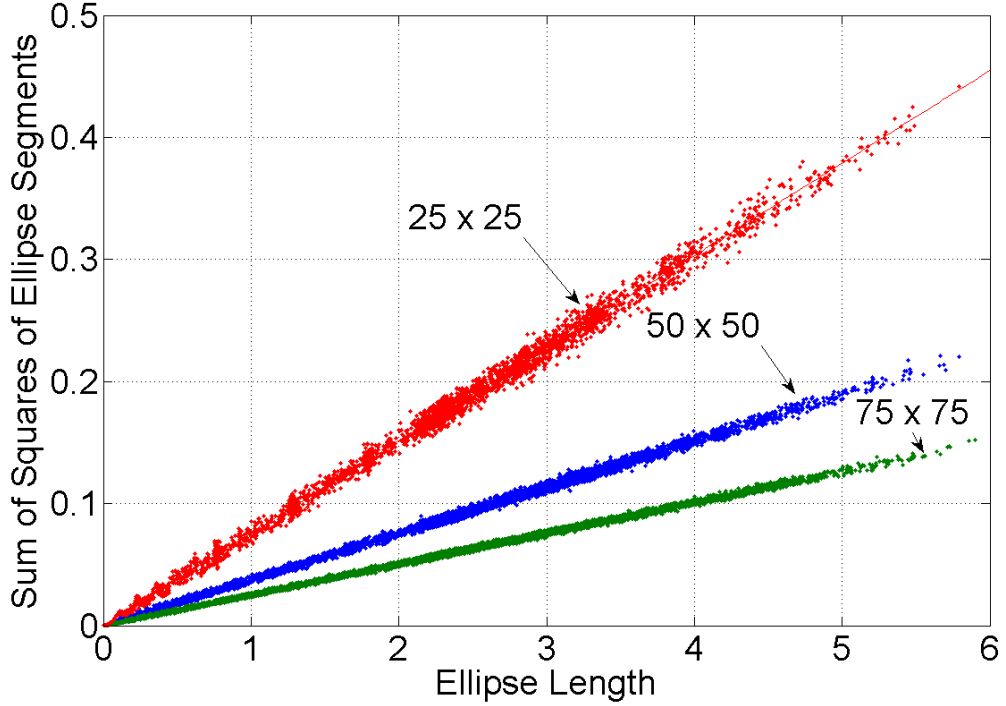


Figure 3.2: Nominal relationships between the sum of squares of a row of \mathbf{A} and the associated ellipse length, for different values of n .

responding row of \mathbf{A} for many ellipses and several pixel dimensions. The linear relationship is a fairly good approximation, and becomes more exact as the resolution of the image increases. Numerical fit of the data shows that $\alpha_1 \approx 0.9469$. If the distribution of segments were independent and uniformly distributed between zero and the diagonal of the square pixel, $\sqrt{2}\delta_x$, then α_1 would be $2\sqrt{2}/3 \approx 0.9428$ in the limit of large ellipses and small δ_x . An estimate of the mean eigenvalue of \mathbf{A} is therefore

$$\bar{\lambda} = \frac{1}{n} \|\mathbf{A}\|_F^2 \approx \frac{1}{n} \sum_{i=1}^m \alpha_1 \delta_x E_i = \frac{\alpha_1}{\beta} \delta_x \bar{E}, \quad (3.29)$$

where \bar{E} is the mean ellipse length.

The average square of the eigenvalues, $\overline{\lambda^2}$, can be estimated by splitting a sum

into two parts:

$$\overline{\lambda^2} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (c^{ij})^2 = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^m (r^{ij})^2 \quad (3.30)$$

$$= \frac{1}{n} \sum_{i=1}^m (r^{ii})^2 + \frac{2}{n} \sum_{i=1}^{m-1} \sum_{j=i+1}^m (r^{ij})^2, \quad (3.31)$$

where r^{ij} is the correlation between the i^{th} and j^{th} rows of \mathbf{A} . The first term can be estimated using the previously-developed approximation for the sum of the squares of a row (equation 3.29):

$$\frac{1}{n} \sum_{i=1}^m (r^{ii})^2 = \frac{1}{n} \sum_{i=1}^m \alpha_1^2 \delta_x^2 E_i^2 = \frac{\alpha_1^2}{\beta} \delta_x^2 \overline{E^2}, \quad (3.32)$$

where $\overline{E^2}$ is the average square of the ellipse length. The second term can be written as

$$\frac{2}{n} \sum_{i=1}^{m-1} \sum_{j=i+1}^m (r^{ij})^2 = \frac{2}{n} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \left(\sum_{k=1}^{N_{ij}} a_{il_k} a_{jm_k} \right)^2, \quad (3.33)$$

where the row cross product as been written in terms of the N_{ij} nominal *intersections* between the i^{th} and j^{th} ellipses, and the values l_k and m_k index only the nonzero components of the inner product. Upon expansion, the square operation yields N_{ij} terms of the form $a_{il_k}^2 a_{jm_k}^2$. Under the previous assumption of uniform distribution, the expected value of this quantity is $\gamma_1 \delta_x^4$ with $\gamma_1 \approx 4/9$. The square operation also yields $N_{ij}(N_{ij} - 1)$ terms of the form $a_{il_k} a_{jm_k} a_{il'_k} a_{jm'_k}$. Treating these quantities similarly, the mean of these terms is $\alpha_3 \delta_x^4$, with $\alpha_3 \approx 1/4$, yielding:

$$\frac{2}{n} \sum_{i=1}^{m-1} \sum_{j=i+1}^m (r^{ij})^2 \approx \frac{2\delta_x^4}{n} \sum_{i=1}^{m-1} \sum_{j=i+1}^m (\gamma_1 - \alpha_3) N_{ij} + \alpha_3 N_{ij}^2 \quad (3.34)$$

$$\approx \frac{2\delta_x^4}{n} \sum_{k=1}^{\sim m^2/2} \alpha_2 N_k + \alpha_3 N_k^2 \quad (3.35)$$

$$= \frac{\delta_x^4 n m^2}{n^2} (\alpha_2 \bar{N} + \alpha_3 \overline{N^2}) \quad (3.36)$$

$$= \frac{\delta_x^2 A}{\beta^2} (\alpha_2 \bar{N} + \alpha_3 \overline{N^2}), \quad (3.37)$$

where the summations over N_{ij} are re-indexed to the $m(m-1)/2 \approx m^2/2$ ellipse pairs, $\alpha_2 = \gamma_1 - \alpha_3 \approx 7/36$, \bar{N} is the average number of nominal intersections between ellipses, $\overline{N^2}$ is the average square of the number of such intersections, and $A = \delta_x^2 n$ is the area of the image of interest. Recombining the terms of equation 3.30 results in

$$\bar{\lambda}^2 \approx \frac{\delta_x^2}{\beta^2} (\beta \alpha_1^2 \overline{E^2} + A(\alpha_2 \bar{N} + \alpha_3 \overline{N^2})), \quad (3.38)$$

giving the increase in eigenvalue power explicitly in terms of the number of intersections.

3.4.2 Bounds in terms of trajectory parameters

The estimates of $\bar{\lambda}$, $\bar{\lambda}^2$, and $\hat{\lambda}_1$ in terms of these physical parameters allow the expression of inversion performance in more informative ways. The absolute bound is simply

$$G_{MIN}(\mathbf{A}) = \frac{1}{\bar{\lambda}} \approx \frac{\beta}{\alpha_1 \delta_x \overline{E}}. \quad (3.39)$$

This result states, for example, that it *may* be feasible to increase the image area A proportionally with the number of pixels, without increasing the size of the ellipses, and without impacting performance. However, for a given area A , which limits the mean size of the ellipses, the image noise power will be inversely proportional to the pixel dimension. The dependence on $\beta = n/m$ means that the noise power could go to zero as the number of measurements m increases.

Using an approximate version of G_P (equation 3.23), the pragmatic bound can

be approximated as

$$G_P(\mathbf{A}) \approx \frac{\beta}{\alpha_1 \delta_x \bar{E} - \alpha_1^2 \delta_x^2 - \frac{1}{\beta} \delta_x^2 A(\alpha_2 \bar{N} + \alpha_3 \bar{N}^2) / \bar{E}^2}. \quad (3.40)$$

Because the second term in the denominator is much smaller than the first, the bound may be expressed as

$$G_P(\mathbf{A}) \approx \frac{\beta}{\alpha_1 \delta_x \bar{E} - \frac{1}{\beta} \delta_x^2 A(\alpha_2 \bar{N} + \alpha_3 \bar{N}^2) / \bar{E}^2}. \quad (3.41)$$

Equation 3.41 is a generalization of the G_{MIN} bound of equation 3.39 which increases the image noise when the mean and mean square of the number of ellipse intersections is significant. The penalty increases with image area and the number of measurements m ; the penalty decreases to zero as the pixel dimension goes to zero.

We are now in a position to summarize a set of design goals for the SAR measurement schemes which should give the best performance.

3.5 Guidelines for trajectory design

The preceding results lead to a simple set of guidelines for the design guidelines for multistatic trajectories of good performance potential.

1. Visit each pixel with equal energy during the measurement process, in accordance with the Welch conditions of equation 3.18.
2. Avoid ellipse intersections; when necessary, ellipses should cross at large angles to minimize their interaction, in accordance with equation 3.41.
3. Use large ellipses when feasible, in accordance with equation 3.39.

Large ellipses generated by bistatic UAV pairs traveling in parallel straight trajectories tend to create ellipses which cross at essentially zero angles from transmit burst to transmit burst. To reduce this tendency, bistatic pairs should follow curved trajectories and travel at different velocities to minimize ellipse intersections.

3.6 Application to example trajectories

We now proceed to apply these guidelines to design and to qualitatively explain the performance of various multistatic trajectories. These examples include deterministic and random trajectories. First, the performance of a variety of trajectory types are surveyed. Second, a more in-depth comparison of a standard deterministic trajectory with a random trajectory is performed.

To enable comparison to previous work, in this section, measurement ellipses are allowed to exceed the bounds of the imaging area. In all but the degenerate case (below), the major axis of the measurement ellipses, for a given transmit burst, increment by the pixel dimension until the ellipse no longer intercepts the image area. This common, but unrealistic, assumption allows simple comparison of various trajectories without the complication of image-edge effects. These effects are dealt with in detail in the final section of this work.

3.6.1 A standard bistatic trajectory

This trajectory is the simplest bistatic trajectory in which two UAVs travel in parallel at equal velocities, with the locus of the center of the vehicles coincident with the bottom edge of the image. This is the trajectory considered for an elliptical integral transform in [66]. This trajectory is not expected to perform well, because it generates many sets of nearly coincident ellipses. The aperture in this case was

± 2.75 image widths with the pitch of ellipse centers equal to $\delta_x = 1/25$ units.

3.6.2 A degenerate case

A degenerate case is easily constructed in which each of n small measurement ellipses (of total length $l_e \approx \pi\delta_x/2$) are designed to fall well within each of the pixel cells. In this case, $m = n$ and \mathbf{A} is proportional to a permutation matrix. This is the *only* case which strictly meets the Welch bound with equality. The image is trivially reconstructed, and the performance matches both G_{MIN} and G_P . Such ellipses can be sampled, for example, by an array of UAVs traveling in parallel at equal heights such that every pixel in the image is situated exactly in-between two UAVs at some point in the measurement. However, such a measurement scheme ignores most of the readily-available data from each transmit burst, using only the smallest sampling ellipses. As a result, this method requires $n = m$ transmit bursts.

3.6.3 A merry-go-round multistatic trajectory

Figure 3.3 shows a candidate multistatic SAR system in which 9 vehicles track each other around a circle above the reflectivity plane to be imaged. All UAVs are at height $h = 1$ image width above the plane. Each UAV generates a set of measurement ellipses with every other UAV (not including itself). The angle of all UAVs on the circle is incremented at 4 degrees per step. After 10 steps (40 degrees) the configuration is identical with the initial configuration, and the measurement process is complete.

Pursuant to the design guidelines, these trajectories are curved. The diameter of the circle traced out by the UAVs gave the best performance at about 0.9 times the image width. The smaller-diameter trajectories are more curved, but gave densely-

sampled inner pixels at the expense of sparsely-sampled outer pixels.

3.6.4 Random trajectories

As a final component of a trajectory survey, we compare a random-location “trajectory” in which sampling ellipses are generated from a swarm of UAVs randomly located above the image of interest. This strategy is analyzed in detail in section 3.7.

There exist a rich set of possibilities for random sampling strategies. The degrees of freedom include (among others): the proportion of UAVs configured as transmitters or receivers; how many of the receivers are responsive to any one transmitter; the spatial distribution of the UAVs; the distribution of velocities (if any single UAV is involved in more than one transmit burst); and the distribution of sample ellipses actually used from a single transmit burst.

For this work, we chose a Monte Carlo scheme in which the locations of a single transmit UAV and a single receive UAV are drawn from a uniform distribution in the volume above the image at heights $h = 1$ to $h = 2$ image widths. This bistatic pair generates and receives only one transmit burst, generating a set of sampling ellipses. This process is repeated many times until a desired number of measurements has been taken.

3.6.5 Performance survey results

Table 3.1 summarizes the results for each of the trajectory types considered. As qualitatively expected, the standard trajectory performed extremely poorly. However, its performance did not approach any of the bounds for G because many eigenvalues were very near to zero. The dominant eigenvalue is poorly estimated in this case, causing (essentially) the failure of the G_p estimate to improve the lower bound.

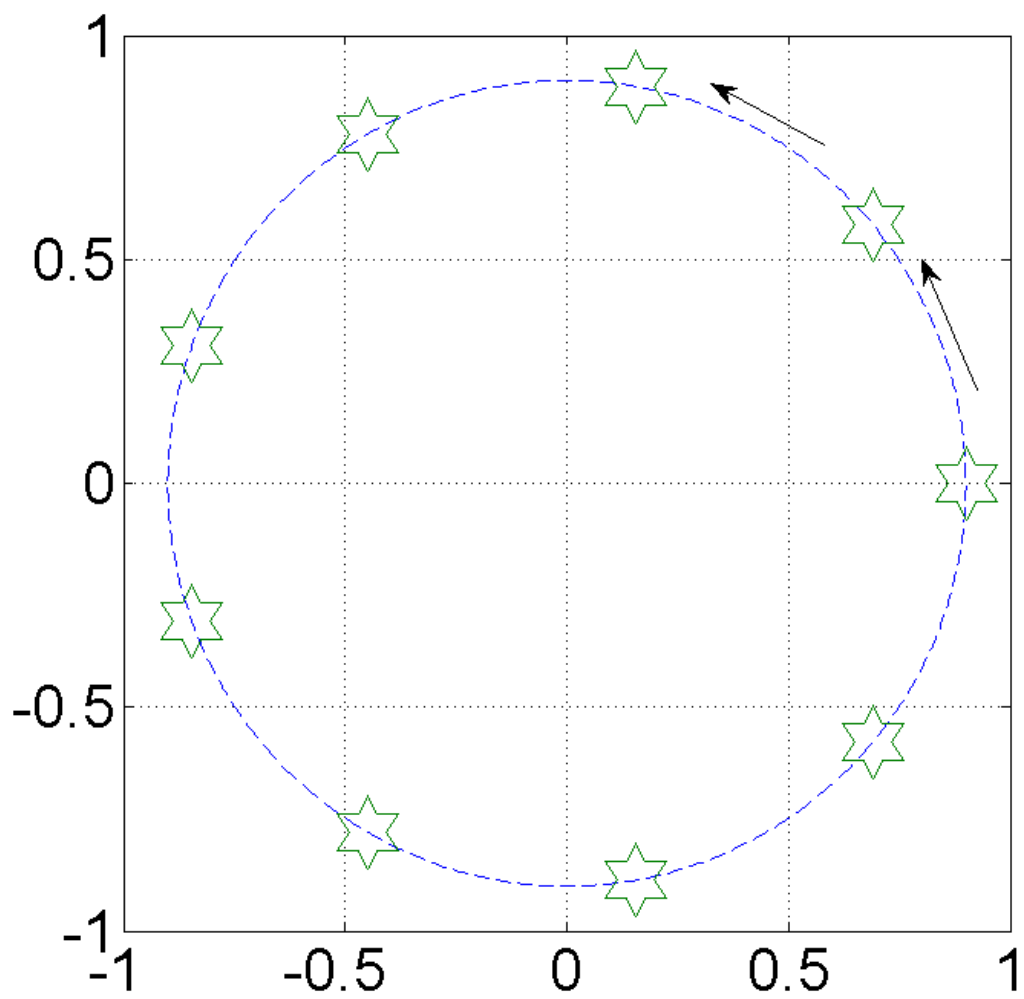


Figure 3.3: A merry-go-round multistatic trajectory with 9 SAR vehicles over a square reflectivity plane.

	Standard	Degenerate	M-G-R	Random
n	2500	2500	2500	2500
m	16348	2500	16838	16088
$\bar{\lambda}$	1.26	4E-3	0.536	0.509
$\hat{\lambda}_1$	22.14	4E-3	39.67	37.84
λ_1	601	4E-3	40.33	38.32
λ_2	577	4E-3	14.60	12.55
G_{MIN}	0.79	253	1.86	1.96
G_p	≈ 0	253	1.99	2.08
Actual G	8461	253	6.72	6.71

Table 3.1: Numerical comparison of SAR performance and bounds

Therefore, this work does not quantitatively capture the details of this trajectory well. We conjecture the reason for this departure is the large number of (essentially) linearly dependent rows of \mathbf{A} in this trajectory. These rows do improve the SNR of one dimension of the matrix, but do little to increase its rank.

In all other cases, the dominant eigenvalue estimate $\hat{\lambda}_1$ is a good approximation to the actual value. The G lower bounds are valid, while the G_p bound gives a better performance estimate.

The deterministic merry-go-round and random sampling cases give essentially the same performance. The ratios of the lower bound G_p to the actual performance are roughly one third in these cases. The subdominant eigenvalue λ_2 is also roughly one third of the dominant eigenvalue λ_1 . The situation suggests that a better lower bound for G could be generated because the addition of $\frac{1}{\lambda_1}$ makes little difference to the sum of $\frac{1}{\lambda_i}$, but directly influences the lower bound. There exist some results in matrix analysis for estimation of the subdominant eigenvalue (see, for example, [64]). However, those results give bounds for strictly positive $\mathbf{A}^T\mathbf{A}$ matrices; in our case, the $\mathbf{A}^T\mathbf{A}$ matrices are still relatively sparse and therefore only nonnegative.

The case matching the Welch bound provides a measure of performance given its small number of measurements, but (taken as a whole) the trajectory is not

impressive. This trajectory uses 2500 transmit bursts, each well-focused on one piece of ground. All other schemes used far fewer bursts; for example, the merry-go-round scheme used only 90 bursts and achieves far better performance with a moderate increase in the number of measurements.

3.6.6 Detailed comparison of random and standard trajectories

We now summarize the performance results of the Monte Carlo simulations and compare these to the standard-trajectory performance. First, we compare in detail the distribution of the first 2500 singular values of \mathbf{A} .

Figure 3.4 highlights a deterministic-looking distribution for the random sampling case, with a sharp cutoff at the low end, exhibiting no singular values below about $\sigma < 0.1$. By contrast, the deterministic version looks rather random; more importantly, it indicates that a large number of singular values exist near the origin.

This small- σ difference shows up even more strikingly in Figure 3.5. The results indicate that random sampling typically outperforms the standard trajectory by 40 dB, a very substantial difference.

To verify the performance improvement due to random sampling methods, we implemented the full measurement and inversion process for a Shepp-Logan test image with measurements corrupted by a Gaussian noise of a constant variance σ_I^2 . Both \mathbf{A} matrices had $m \approx 9000$. Figure 3.6 shows a side-by-side comparison the resulting images.

The random-sampling inversion is visually indistinguishable from a perfect image, while the standard-trajectory inversion, while recognizable, is clearly inferior. The distortions in the standard image appear striped because the 2D eigenfunctions

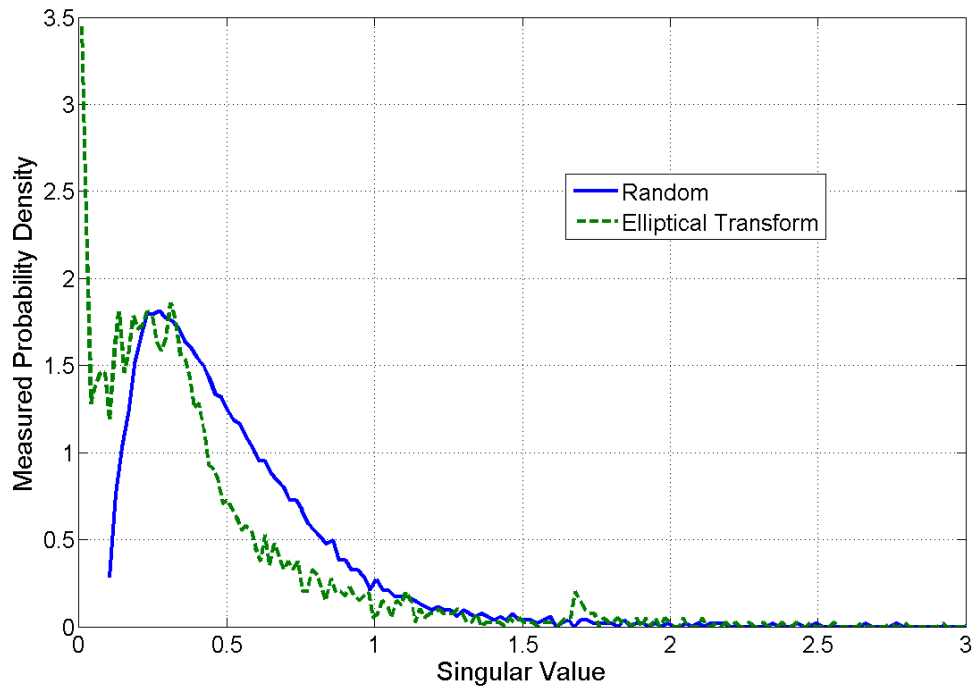


Figure 3.4: The empirical singular value distributions of \mathbf{A} for \mathcal{R}_{ef} and for a random sampling trial, with $m \approx 12000$.

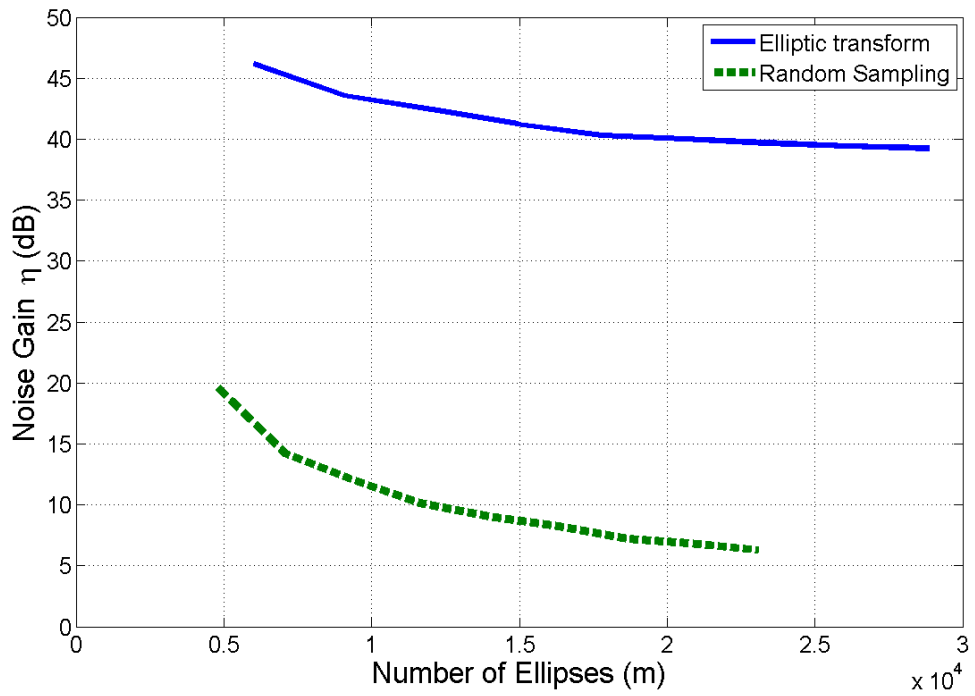


Figure 3.5: Noise gain of deterministic and random sampling as a function of the measurement size m and constant $n = 50 \times 50$.

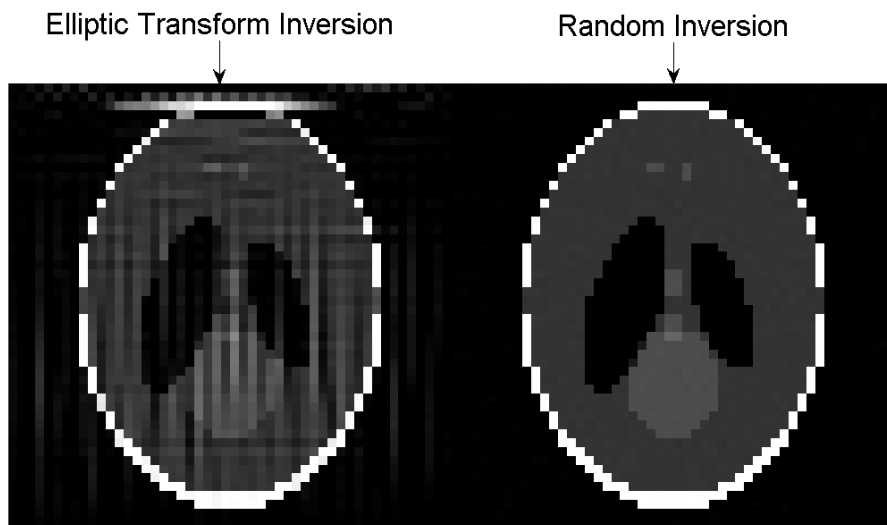


Figure 3.6: Comparison of reconstructed images with the same measurement noise power, using $m \approx 9000$.

associated with the smallest singular values exhibit these types of shapes. Therefore, random sampling produces stable inversions, and in fact can easily outperform the inversion of standard deterministic trajectories.

The relatively excellent performance of the random trajectory motivates further study into its performance.

3.7 Performance estimation by random matrix theory

The non-zero elements of the sparse matrix \mathbf{A} are the arc lengths of randomly placed ellipses cut by a rectangular grid. We now examine the consequences of treating the elements A_{ij} as random variables with marginal density function

$$p_x(x) = (1 - s)\delta(x) + sP_x(x) \tag{3.42}$$

where s is the sparsity of \mathbf{A} . For dense matrices ($s \approx 1$) with i.i.d. A_{ij} entries, large m and n with $\beta = n/m$, the eigenvalues of $\mathbf{A}^T \mathbf{A}$ are distributed according to the Marchenko-Pastur (MP) asymptotic law [67]:

$$p_\lambda(\lambda) = \frac{1}{2\pi\beta m\sigma_A^2\lambda} \sqrt{(b - \lambda)(\lambda - a)} \tag{3.43}$$

where $\lambda > a = m\sigma_A^2(1 - \sqrt{\beta})^2$, $\lambda < b = m\sigma_A^2(1 + \sqrt{\beta})^2$, and σ_A^2 is the variance of A_{ij} . The general problem, with jointly distributed A_{ij} entries and $s \ll 1$, is open [68]. Therefore, the MP distribution must be approximate for highly sparse matrices such as \mathbf{A} . The MP parameter $\sigma_A^2 = \|\mathbf{A}\|_F^2 / (nm) - \mu_A^2$, where μ_A is the mean of A_{ij} , can be calculated directly from the matrix. The appropriate version of the G

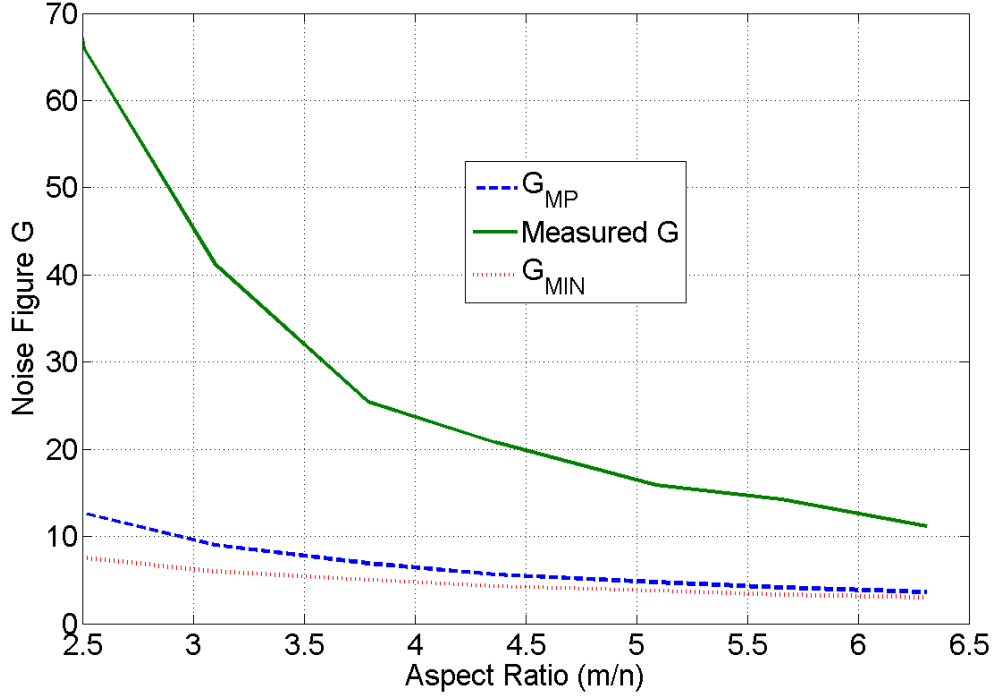


Figure 3.7: Comparison of G estimates and actual performance for an $n = 75 \times 75$ image using random trajectories and various values of m .

figure of merit (equation 3.10) for continuous distributions is simply the expectation of $1/\lambda$:

$$G_{MP} = E\left[\frac{1}{\lambda}\right] = \int_a^b \frac{1}{\lambda} p_\lambda(\lambda) d\lambda = \frac{1}{\sigma_A^2(m-n)}. \quad (3.44)$$

In preliminary work [69], we raised the conjecture that G_{MP} may be an upper bound for the actual performance. The surmise was based on observations (in common with Nagao [68]) that the eigenvalue distribution variations due to matrix sparsity were most prominent in the high end of the spectrum. However, the conjecture cannot be maintained in the light of further numerical experience, in which we have found significant deviations also in the low end of the spectrum. Therefore, G_{MP} should be taken as an estimate and not a bound.

Figure 3.7 shows the measured performance of G compared to G_{MIN} and G_{MP}

for random trajectories of higher resolution $n = 75 \times 75$. The parameters for the G estimates come from direct measurements of the associated \mathbf{A} matrix. In this case, for larger m , the estimates exceed the actual performance by a consistent factor of three. Thus, G_{MP} gives a reasonable order-of-magnitude estimate of G . Better estimates could come from better approximations for $p_\lambda(\lambda)$.

Figure 3.7 also suggests that G_{MIN} and G_p will converge at large m . Indeed,

$$\frac{G_{MP}}{G_{MIN}} = \frac{m(\sigma_A^2 + \mu_A^2)}{\sigma_A^2(m - n)} = \left(\frac{m}{m - n} \right) \left(\frac{\sigma_A^2 + \mu_A^2}{\sigma_A^2} \right). \quad (3.45)$$

The first factor clearly decreases to unity as m increases. Furthermore, the second factor also approaches unity because \mathbf{A} is sparse, with sparsity inversely proportional to \sqrt{n} . As n increases, the term μ_A^2 becomes negligible with respect to σ_A^2 . Therefore, for large enough problems, and to the extent that G_{MP} is a good estimate of the actual performance, random sampling asymptotically reaches the best possible performance of *any* sampling strategy of equal \mathbf{A} energy.

3.7.1 System scaling equations

Finally, equation 3.44 can be recast in terms of physical parameters of interest in the SAR system. It can be shown by simple counting arguments that the sparsity s of \mathbf{A} is inversely proportional to \sqrt{n} . Therefore, the variance of A_{ij} is proportional to δ_x^2/\sqrt{n} , where δ_x^2 is the area corresponding to a pixel in the reflectivity plane. But $\delta_x^2 = A/n$, where A is the area of the image. Therefore,

$$G_{MP} = \kappa \frac{n^{3/2}}{A(m - n)} \quad (3.46)$$

for some constant κ . Recall that the measurement noise has been modeled as i.i.d. Gaussian, which is approximately correct in the case where the SAR system noise is dominated by the electronic noise in the receiver amplifiers. Under this model, improvements in linear resolution on the same area require both the electronic noise bandwidth and the input noise variance σ_I^2 to increase proportionally with \sqrt{n} . Therefore the image noise variance can be expressed as

$$\sigma_O^2 = \kappa' \frac{n^2}{A(m-n)}. \quad (3.47)$$

Equations of this type are useful as scaling rules for system design. For example, consider a working SAR system in which $n = 10,000$ and which can process $m = 30,000$ measurements. According to equation 3.47, a new system needing twice the pixels over the same area will need to boost transmit power by 8 times (9 dB) to maintain the pixel quality of the old product. Alternatively, m could be increased to 100,000 without modification to the transmit power.

3.8 Edge effects and variable resolution imaging

In all simulations up to this point, sampling ellipses have been allowed to exceed the image bounds. This common but unrealistic supposition presumes that the contributions to the measurement from outside the image bounds are zero, or are otherwise known. Most versions of SAR systems limit this type of problem by designing the antenna system to spatially filter out reflections which occur outside the area of interest. In a final application, we consider what may be done to retain an omnidirectional-style antenna strategy in a multistatic system.

After deletion of all sampling ellipses which exceed the image bound, those re-

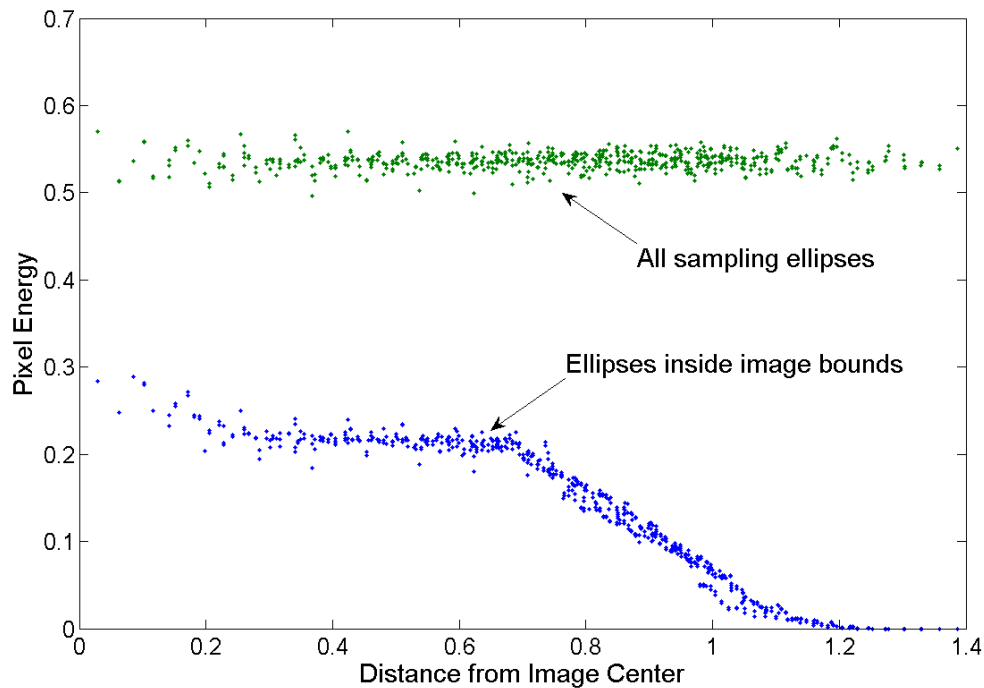


Figure 3.8: The effect on pixel energy of deleting out-of-bounds sampling ellipses in an $n = 50 \times 50$ merry-go-round trajectory. (An image width = 2 distance units).

maintaining ellipses involving pixels near the image edges will have portions which are generally parallel with the image edge. In addition, the sampling density near the edges will be significantly decreased because the possible loci of ellipse centers decreases to (generally speaking) an area normal to the image-edge contour. The situation for the merry-go-round trajectory is shown in Figure 3.8.

The pixel energy is the sum of the squares of the column of \mathbf{A} associated with that pixel. Plotting these energies versus the distance from the center pixels shows that the original scheme excited each pixel approximately equally. Removal of the out-of-bounds ellipses significantly decreases the pixel energy everywhere because the number of sampling ellipses decreases from $m = 16838$ to $m = 4594$. The pixels distant from the center are most affected. Because some pixels far from the center

are rarely sampled, the inversion performance of this \mathbf{A} is expected to be poor, and in fact, $G = \infty$ in this case because many pixels in the corner of the image area are never sampled at all.

By the first design guideline, we seek an after-the-fact method to equalize the energy in the columns of the measurement matrix \mathbf{A} . The basic idea proposed here is to combine the energy of adjacent low-energy pixels into superpixels, but only where needed to meet some expectation of pixel energy. This expectation would typically be something close to the energy of the central pixels. This technique results in a variable-resolution scheme in which the reconstructed image will have reduced resolution near the image boundary.

The technique is efficiently described by replacing the n pixels \mathbf{f} with a set of $n' \leq n$ pixels \mathbf{f}' , related by

$$\mathbf{f} = \mathbf{P}\mathbf{f}', \quad (3.48)$$

where the $n \times n'$ sparse matrix \mathbf{P} defines the expansion of the superpixels in \mathbf{f}' . Rows of \mathbf{P} contain as many ones as there are original pixels in each of the new pixel set. Therefore, the matrix \mathbf{P} is a non-square generalization of a permutation matrix, which contains n non-zero elements. Other intuitions and design methods for \mathbf{P} are possible, which has the form of a pre-conditioning matrix.

The new matrix equations may be written as

$$\mathbf{A}(\mathbf{P}\mathbf{f}') = \mathbf{r} \quad (3.49)$$

$$\mathbf{A}'\mathbf{f}' = \mathbf{r}, \quad (3.50)$$

where it is hoped that $\mathbf{A}' = \mathbf{A}\mathbf{P}$ now has improved performance. Note that the measurements \mathbf{r} are unaffected by this technique. The new system in equation 3.50 can be solved as usual for \mathbf{f}' , and the original pixels can be recovered by $\mathbf{f} = \mathbf{P}\mathbf{f}'$ as



Figure 3.9: Superpixel pattern of dimensions 8, 4, 2, and 1, from light to dark (left). Noiseless reconstruction with no reflectors in superpixel locations (middle). Noiseless reconstruction with bright single-pixel reflector within the marked superpixel location (right).

necessary.

Because the merry-go-round trajectory is symmetric and regular, it is straightforward to construct a superpixel pattern which will significantly increase the effective minimum pixel energy. Figure 3.9 shows a superpixel pattern which increases the effective minimum pixel energy from zero to about 0.09 units (compare Figure 3.8). Thus, roughly one third of the original worst-energy pixels are swept up into superpixels. The measured inversion performance using this superpixel pattern is $G = 26.8$, that is, about four times worse than the performance of the same trajectory without deleting the out-of-bounds ellipses, while using about 4 times fewer sampling ellipses. If the number of measurements is increased to $m = 33767$ (about double that of the original simulation), then the inversion performance significantly improves to $G = 1.25$ (or, about 5 times better than that of the original simulation). Therefore, the variable-resolution technique has indeed enabled the limited-ellipse inversion performance, as measured by G , to bracket that of the original scheme.

This simple technique is not without its drawbacks. The middle diagram of Figure 3.9 shows essentially perfect reconstruction if there do not happen to be re-

flectors in the superpixel area. However, the overall performance of the inversion can be degraded when significant variation in reflectivity exists within a superpixel. This performance degradation is due to a type of aliasing: sampling ellipses passing through a superpixel may see significantly varying reflections depending on the location of the ellipse segment within the superpixel. The effect is demonstrated in the right side of Figure 3.9, where a single bright pixel is introduced into the measurement vector \mathbf{r} at the marked location. The effect of this reflector is to introduce visible distortions in the reconstructed image which can extend throughout a large portion of the image. If there are many such superpixels, then the image distortions due to aliasing can be very large.

Therefore, a classic tradeoff exists between improving the measurement-noise performance by the use of variable resolution, and the introduction of aliasing distortions due to too many superpixels with significant reflectivity variation. Depending on system parameters, further algorithm development may be necessary to relieve this tradeoff.

3.9 Random-location sampling and compressed sensing

Because the random-sampling aspects of the present work are akin to other sampling concepts currently under development in the literature, it is appropriate to provide a context in which these ideas may be compared and contrasted. In the related area of *compressed sensing* (CS) [53, 54], an image, which is sparsely represented in some basis, can be reconstructed from projections of the image onto purely randomly-constructed basis functions. The signature feature of CS reconstruction is that the number of projections m required to faithfully reconstruct a sparse image can be

substantially *less* than the number of pixels n . The now-standard method of *basis pursuit* [70], for example, can reconstruct a sparse image by finding that \mathbf{x} with minimum ℓ_1 norm which is consistent with the measurement information $\mathbf{Ax} = \mathbf{r}$.

In CS, \mathbf{A} is called an *information operator* and is typically assumed to be dense and possess independent, identically-distributed (i.i.d.) entries; Donoho proves (using the MP law, which is valid for very large matrices of this type) that information operators so constructed are very likely to be suitable for compressed sensing [54]. In contrast, under our SAR measurement framework, what we call “random sampling” describes a physical line integral over an ellipse with a random location within the image area. As noted previously, the resulting operator \mathbf{A} is neither i.i.d. nor dense.

To demonstrate whether the advantages of CS survive SAR’s physical sampling constraints in a good trajectory, we follow [71] in which Shepp-Logan phantom images are reconstructed by minimizing the gradient or total variation (TV) in an image under the information-operator constraint $\mathbf{Ax} = \mathbf{r}$. The Shepp-Logan and other artificial piecewise-constant phantoms are particularly well suited (perhaps, too well suited) to this technique because their gradients are even more sparse than the image itself. We used the software package “ ℓ_1 -magic” [72] to generate a minimum-TV reconstruction from a random-location SAR \mathbf{A} matrix and its projections \mathbf{r} .

Figure 3.10 shows the results for a 75×75 image, using 25 randomly-located bistatic transmit events, generating a total of 1765 measurements; here, \mathbf{A} is a 1765×5625 matrix. For comparison purposes, the minimum- ℓ_2 reconstruction from the standard pseudoinverse is shown. The minimum-TV reconstruction is visually indistinguishable from the original image, indicating that CS concepts can remain in force despite the sparsity and correlations of the \mathbf{A} matrix, enforced by the physical constraints of the SAR system.

Naturally, CS works best without measurement noise, which can markedly erode

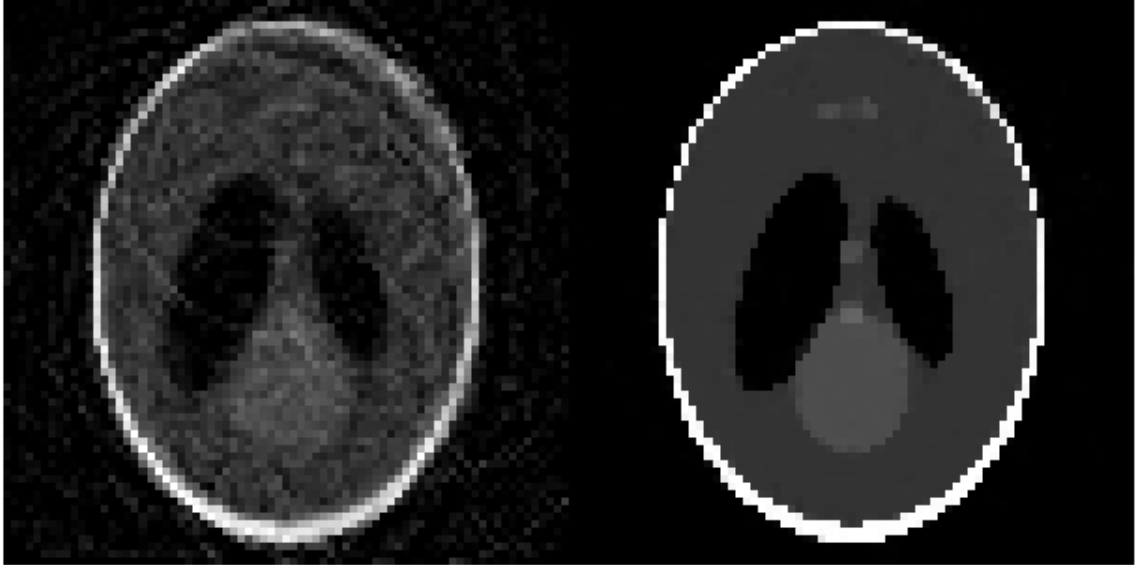


Figure 3.10: Noiseless minimum- ℓ_2 norm reconstruction (left) and minimum-TV reconstruction (right) for a $75 \times 75 = 5625$ pixel image and $m = 1765$ random-location SAR measurements.

its gains [73]. However, alternative methods for sparse reconstruction in the presence of noise is a current topic of research, e.g. [74]. Such research will undoubtedly be applicable for numerical solutions to the noisy SAR problem. However, the current paucity of analytic performance tools for CS-related reconstruction methods (such as those tools we used from linear algebra for the maximum-likelihood methods) will make it more difficult to develop trajectory guidelines for optimum CS performance in the presence of noise.

3.10 Conclusions

We have developed several bounds for the image-noise performance of a multistatic SAR imaging system, based on easily-measured statistics of the measurement matrix \mathbf{A} . The Welch bound and the development of a physical interpretation of the matrix analysis results lead to a set of guidelines for the design of UAV trajectories. The

results of several example UAV trajectory designs were explained using this approach.

We found that random sampling strategies can perform as well as the best deterministic method found in this work, and gave an approximate argument based on random matrix theory that this in fact should be the case for large problems. This result is reminiscent of similar-sounding results in information and coding theory. Depending on system considerations (including the cost of UAV coordination and station-keeping, and constraints on the number of UAVs), one or the other strategy may be preferable.

Future extensions of this work might include better handling of image-edge effects; the development of better bounds using estimates of subdominant eigenvalues; the extension of the figures of merit to include the number of transmit bursts required; the development of tighter pragmatic upper bounds; and the development of more specific methods to design deterministic UAV trajectories; and the improvement of G_{MP} by better estimation of the distribution of eigenvalues in a sparse positive random matrix.

Chapter 4

Simplified-hardware ranging systems using wavelet decomposition

Previous chapters of this work have developed the signal processing consequences of reconstructing an image from simple, down-looking UAVs operating in general three dimensional orbits. The price of simplicity at the UAV is higher computational burden at the central processor. While a main goal of this work is to examine the system impacts of this general strategy, leaving *some* processing on the UAV can be advantageous. For example, if the UAV is able recognize synchronization pulses (such as might come as "side information" from the direct, non-reflected query from a transmitter), the UAV can reduce the amount of information it must send to the central processor. In this chapter, we develop families of phased-coded waveforms can be decomposed at the receiver in an extremely simple manner, yet the decomposition yields good pulse compression.

Phase-coded waveforms with good autocorrelation properties have long been

studied as a means to provide both high range resolution and good signal-to-noise properties. Such sequences were generally characterized by their length and by some statistic of the system's sidelobe performance from a matched filter receiver. More recently, ranging systems have been studied in which the transmit waveform is substantially modified in order to gain excellent sidelobe performance if all targets are known to be within a finite tolerance of a given range. These systems are based on the *zero autocorrelation zone* (ZAZ) sequences [75, 76, 77, 78, 79, 80]. In ZAZ systems, the transmit waveform modifications are simply cyclic extensions of the receiver filter sequence.

In the present work, we propose two novel waveform modification techniques which not only retain excellent sidelobe performance, but also allow significant simplification of the receiver hardware. To achieve this simplification, we shall find it convenient to codify and to extend previous research by casting the receive hardware implementation in the language of a wavelet transform.

There exists a fundamental connection between wavelet transforms and problems in radar and sonar. Optimal range and velocity detection are implemented by an appropriate wavelet transform because such transforms implement optimal correlation processing on translations (range) and scales (velocity) [81, 82, 83, 84]. For example, [85, 86] propose the idea of directly using a mother wavelet function ψ as the transmit waveform, and then computing the wavelet transform of the received signal to decompose the signal into its frame coefficients. Neglecting modulation and equalization effects, the frame coefficients of a point-reflector receive waveform are all zero except for one, *so long as* the received waveform is on-grid of scale and translation.

This last condition is an inherent issue for the direct use of wavelets as transmit waveforms, because only in unusual situations would potential targets be situated

exactly on any grid of range and velocity. The problem arises from the shift-variance property of both continuous and discrete wavelet transforms, that is, the frame coefficients are (generally) not just shifted in time when the input waveform is shifted in time. In recent work, Sveinsson [87] addresses this issue by developing *almost translation invariant* wavelet transforms for improvement of SAR images.

Modern problems in the disparate fields of sequence design and wavelet theory are tightly related because of their common ancestry in the fundamental work of Welti [88], Golay [89] and others [90, 91]. The first part of this work is to review sequence design and matched filter implementations for *E*-sequences [90], which, we show, can be very simply posed and solved in the language of the DWT. We compare the resulting DWT implementation to similar implementations derived by other methods [92, 93, 94]. A DWT-decomposable set of sequences is developed based on the original, half-century-old research, which nonetheless possess many characteristics of current research interest. The construction yields ZAZ sequences; the construction also produces zero-correlation-zone (ZCZ), loosely-synchronous, and mutually-orthogonal-complementary sets [75, 95, 96, 79, 97, 98].

The ZAZ concept achieves good autocorrelation, for sequences with good *periodic* correlation up to an index τ , by a cyclic extension of length $\pm\tau$ of the input sequence. We generalize the ZAZ concept by explicitly constructing the sequence extensions necessary to achieve perfect autocorrelation for any given sequence. We use this construction to extend the zero correlation zone of the DWT-decomposable sequences beyond the previously-studied, cyclic limits. The construction opens up new tradeoffs between transmit power, sequence length, and receiver complexity.

The contributions of this work can be grouped in three main areas. First, we employ the concepts and tools of the DWT to encapsulate and extend a body of previous research on efficiently-analyzable waveforms. Second, we propose a general

sequence extension technique which extends the zero correlation zone width of any sequence; in particular, we show significant improvements in the ZAZ width for particular sets of sequences which are DWT-decomposable. Finally, we propose a specific transmit waveform construction which allows an additional factor-of-two reduction in the receiver hardware complexity.

The chapter is organized as follows. In section 4.1, we review the DWT, and propose a set of example sequences which are DWT-decomposable and have good autocorrelation properties. In section 4.2, we propose a general sequence extension method which widens the zero correlation zone past its cyclic limits. The method is applied to our sequence set. Section 4.3 develops a transmit-sequence construction in which all outputs of the DWT decomposition are efficiently used, thereby gaining a factor of two in implementation complexity. Finally, section 4.4 contains concluding remarks.

4.1 Background

The forward discrete wavelet transform (DWT), as considered here, takes a one-dimensional signal and generates a two-dimensional signal with samples at half the original rate. The components of the DWT output are downsampled from analysis filters $C(z)$ and $D(z)$ (see Figure 4.1). Not all analysis filters C and D admit a reversible transformation. Sufficient conditions for C and D to represent finite impulse response (FIR) filters, and also to produce outputs from which the original sequence can be reconstructed perfectly, are as follows [99]. First, the aperiodic autocorrelation of the sequence defining the filter C is constrained to be zero at non-zero even shifts:

$$\sum_{k=0} c_k c_{k+2m} = 0 \quad (4.1)$$

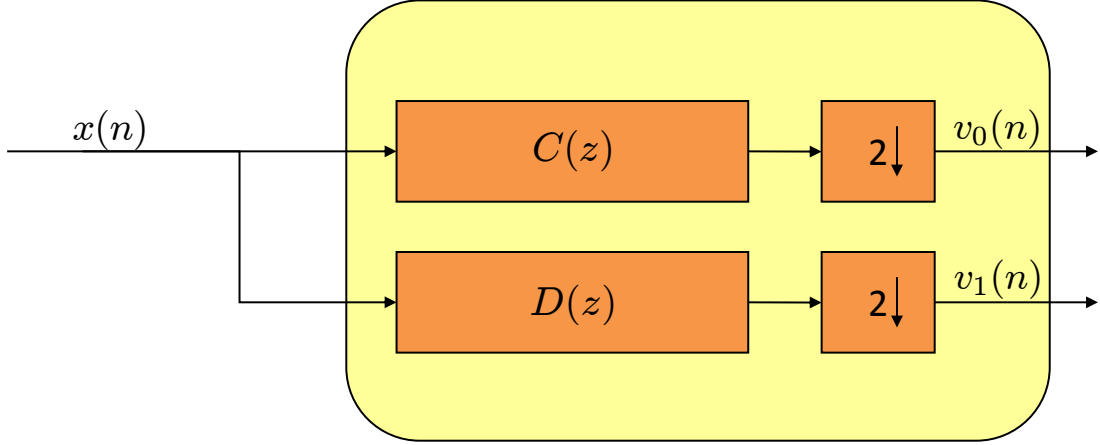


Figure 4.1: A forward discrete wavelet transform.

for all integer $m \neq 0$, and where c_k represents the k^{th} tap weight of the sequence \mathbf{c} defining the transfer function $C(z)$, $k \geq 0$. In classical terms, sequences \mathbf{c} obeying the constraint in equation 4.1 are called *E-sequences* [90], which can easily generate complementary pairs of sequences [89]. In modern parlance, equation 4.1 is a necessary condition for the orthogonality of continuous wavelet functions generated from \mathbf{c} using dilation methods [100, 101].

Furthermore, the filter D is generated from C by the relationship:

$$d_k = (-1)^k c_{L-k-1}, 0 \leq k < L. \quad (4.2)$$

where k spans the length of the nonzero values comprising \mathbf{c} . Again depending on era, the sequences \mathbf{c} and \mathbf{d} are then known as *mates* [88], or as the other's *quadrature matched filter* (QMF).

The preferred implementation of the DWT recasts the filtering problem using the polyphase decomposition technique, in which the input signals and the C and D

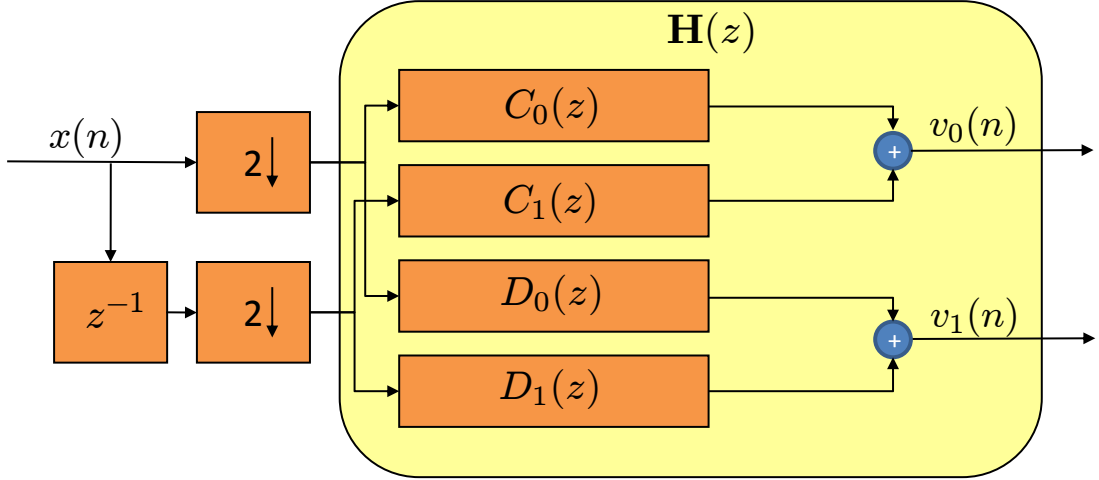


Figure 4.2: A polyphase implementation of a discrete wavelet transform.

filters are split into their respective odd and even interleaves. The resulting structure is a 2×2 linear filter bank which operates at half the input sample rate (see Figure 4.2). The input/output relationship can be written in matrix form:

$$\begin{bmatrix} V_A(z) \\ V_B(z) \end{bmatrix} = \begin{bmatrix} C_0(z) & C_1(z) \\ D_0(z) & D_1(z) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} \quad (4.3)$$

$$= \mathbf{H}(z) \begin{bmatrix} X_0(z) \\ z^{-1}X_1(z) \end{bmatrix} \quad (4.4)$$

where the 0 subscript denotes the even interleave, and the 1 subscript denotes the odd interleave, of their respective antecedents. The input processing prior to the \mathbf{H} operation is simply a deserialization of the input stream in which the upper sample $x_0(n)$ is newer than the lower sample $x_1(n)$. When C and D obey equations 4.1 and 4.2, then \mathbf{H} is a paraunitary matrix. Such matrices are nicely factorizable [102], providing an elegant means to generate efficient implementations.

A typical use of a DWT decomposition is to implement a multiresolution analysis by simply applying the same DWT to either or both of the output sequences v_A and v_B . In this work, we simply view the decomposition as a very efficient means to implement the matched filters C or D .

The DWT implementations shown in Figures 4.1 and 4.2 are functionally identical. In the present work, the direct-form DWT provides functional intuition for the use of the DWT, while the polyphase version produces elegant correlation-filter implementations which are suitable for high speed applications because the resulting filters are natively pipelined to operate at half rate.

4.1.1 Sequence constructions

There are infinitely many E -sequences, each giving a feasible DWT decomposition. The usual intuition provided for the design of wavelets is to define a lowpass function using \mathbf{c} ; in such a case, its mate \mathbf{d} is a highpass function. However in ranging matched-filter applications, it is advantageous to design waveforms (and their matched filters) which approximate a flat allpass response over all frequencies, such that the matched filter output better approximates a delta function. Design methods for \mathbf{c} can vary greatly; for example, we showed in [103] that suitable binary sequences can be generated by enforcing a characteristic-phase constraint (borrowed from m -sequence theory) together with the E -sequence constraint.

In order to demonstrate the concepts of this work, we desire sequences which produce exceptionally simple DWT decompositions *and* have good correlation properties under the transmit waveform modifications which we shall propose. For this purpose, we generate candidate E sequences by the vectorization of a construction

method originally given by Welti:

$$\mathbf{W}_{l+1} = \begin{bmatrix} \mathbf{A}_l & \mathbf{B}_l & \mathbf{A}_l & -\mathbf{B}_l \\ \mathbf{A}_{l+1} & \mathbf{B}_{l+1} & & \end{bmatrix} = \begin{bmatrix} \mathbf{A}_l & \mathbf{B}_l & \mathbf{A}_l & -\mathbf{B}_l \\ \mathbf{A}_l & \mathbf{B}_l & -\mathbf{A}_l & \mathbf{B}_l \end{bmatrix}, \quad (4.5)$$

where \mathbf{A}_l and \mathbf{B}_l have an equal number of columns and split the square matrix \mathbf{W}_l in half. If all rows of a matrix \mathbf{W}_l are E sequences, then the rows of \mathbf{W}_{l+1} are also E sequences. Starting with a degenerate E sequence (the delta function $\mathbf{W}_0 = 1$), we have $\mathbf{A}_0 = 1$, and \mathbf{B}_0 is the empty matrix. Then a sequence of E sequence sets \mathbf{W}_l starts with:

$$\mathbf{W}_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \mathbf{W}_2 = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \quad (4.6)$$

Despite the construction's antiquity, \mathbf{W} has many characteristics under current study in the area of zero correlation zone (ZCZ) or loosely-synchronous (LS) sequence sets.

1. \mathbf{W} is a Hadamard matrix.
2. Every row \mathbf{w}_j of \mathbf{W}_l is a length $L = 2^l$ E sequence.
3. Every row \mathbf{w}_j is a zero autocorrelation zone (ZAZ) sequence, exhibiting perfect periodic autocorrelation up to an index of $\pm L/4$.
4. \mathbf{W} contains the mate and reversed sequence of every row \mathbf{w}_j , within a sign.
5. The rows \mathbf{w}_j and $\mathbf{w}_{j+L/2}$ are mates and complementary pairs [89].
6. The rows \mathbf{w}_j and \mathbf{w}_{j+1} are complementary pairs and also form a ZCZ set with

parameters $\text{ZCZ}(L, 2, L/4)$ [75]. This pair matches the ZCZ bound for binary sequences discussed in [76].

7. The set of all rows of \mathbf{W} is a complementary sequence set [97].
8. The top and bottom halves of \mathbf{W} form mutually orthogonal Golay complementary sets (MOGCS) [104].

4.1.2 DWT implementations

Any two-dimensional degree- K paraunitary filter \mathbf{H} can be factored in the form:

$$\mathbf{H}(z) = \alpha \left\{ \prod_{k=K}^1 \mathbf{R}(\theta_k) \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \right\} \mathbf{R}(\theta_0) \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix} \quad (4.7)$$

for some scaling α , and where $\mathbf{R}(\theta)$ is any scaled 2×2 rotation matrix:

$$\mathbf{R}(\theta) = \begin{cases} \begin{pmatrix} 0 & \sin \theta \\ -\sin \theta & 0 \end{pmatrix} & \text{if } \theta = \pm \frac{\pi}{2}, \\ \begin{pmatrix} 1 & \tan \theta \\ -\tan \theta & 1 \end{pmatrix} & \text{otherwise.} \end{cases} \quad (4.8)$$

Applying the planar factorization method described in [102, pp. 727-732] to rows of \mathbf{W}_l as the \mathbf{c} sequence yields factorizations of the form

$$\mathbf{H}_l(z) = \left\{ \prod_{j=l-1}^1 \mathbf{R}(b_j \frac{\pi}{4}) \begin{bmatrix} 1 & 0 \\ 0 & z^{-2^{j-1}} \end{bmatrix} \right\} \mathbf{R}(b_0 \frac{\pi}{4}) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (4.9)$$

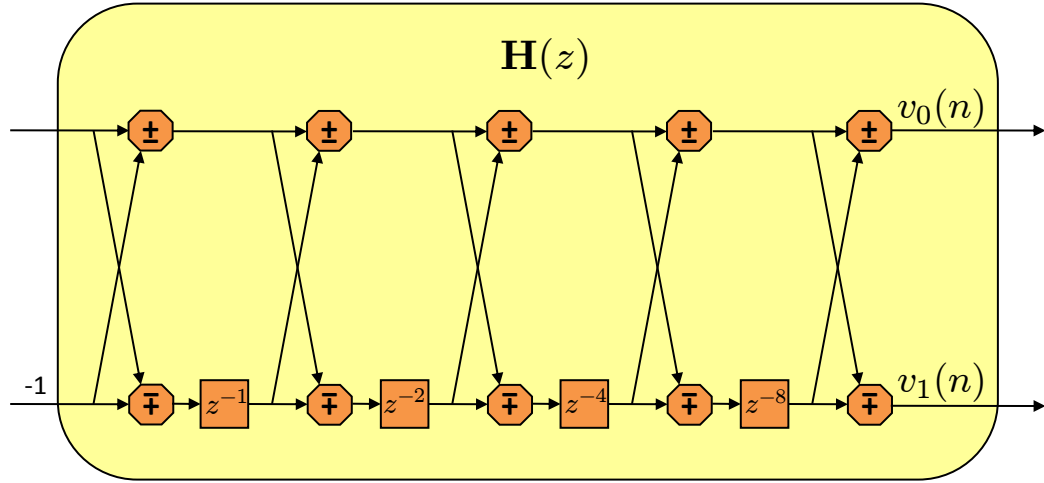


Figure 4.3: DWT decomposition hardware for \mathbf{H} with $L = 32$.

where $b_j \in \{\pm 1\}$. The 2^l possible sets of b values corresponds to the 2^l rows of \mathbf{W}_l . A general circuit implementing equation 4.9 is shown in Figure 4.3. In vertical pairs, corresponding to a $\mathbf{R}(\pm\frac{\pi}{4})$ multiplication, one of the combinational blocks is an addition and one block is a subtraction, depending on the corresponding value of b_j . When the operation is subtraction, the subtrahend always comes from the signal depicted by the oblique arrows. The simplicity of this implementation arises from the fact that about half of the θ_k in equation 4.7 are zero, leading to identity matrices in $\mathbf{R}(0)$ which can be deleted.

Because implementation complexity of additions and subtractions are nearly identical, we shall count a subtraction as an addition. For a corresponding \mathbf{c} sequence of length L , the circuit requires $2 \log_2 L$ additions, no multiplications, and $L/2 - 1$ delay elements. For the price of some exclusive-OR gates, adding programmability to the set of b_j polarities, this circuit can be made to act as a component of a matched filter for any sequence in \mathbf{W} . The necessary detail of time-reversed sequences (or

inverted filter weights) to correctly implement the correlation does not change the form of the implementation because \mathbf{W} contains the time-reversed versions of each of its sequences.

This implementation form looks very like implementations developed in [88], and in generalizations developed in later decades [92, 93, 105, 106, 94]. Aside from its easy derivation using the tools of the DWT, the DWT version's main advantage is that it directly produces an efficient circuit which runs at half the usual rate. Such a circuit is therefore very suitable for high-speed implementation.

With reference to Figure 4.1, the DWT directly produces only every-other value of the desired matched filter $C(z)$, and in addition it produces a seemingly-useless output $D(z)$. The remainder of this correspondence is devoted to the design of transmit waveforms for which the use of this simple DWT decomposition at the receiver is appropriate. In so doing, we shall introduce an extension of the ZAZ concept.

4.2 Generalized zero autocorrelation zone extensions

In OFDM communications [41], transmit waveforms are cyclically extended by the span of the expected intersymbol interference present in the transmit channel. These extensions enable the receiver to implement a circular (periodic) deconvolution even though physical channel's operation is modeled as a linear convolution. Essentially the same idea is applied for ranging systems and ZAZ sequences, in which a transmit sequence \mathbf{c} is cyclically extended such that, for a finite range of time, the nominal output of the matched filter is the same as the periodic autocorrelation of \mathbf{c} . Thus, if \mathbf{c} exhibits perfect periodic autocorrelation within a finite range of offsets, then a

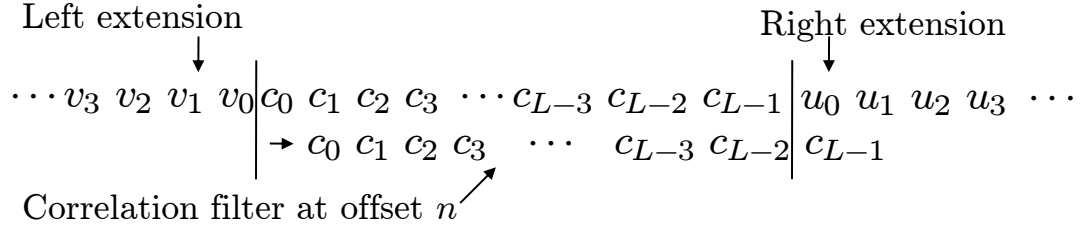


Figure 4.4: Extending the sequence \mathbf{c} for perfect correlation filter outputs.

point reflector seen through the matched filter exhibits perfect sidelobes so long as it is agreed that any such reflector be within a corresponding finite distance range.

However, we may recognize that the extension need not be cyclic to retain the advantages of ZAZ sequences, for ranging applications. For example, the pulse sequence $\mathbf{c} = \{1, 1, 1, 1\}$ gives an unimpressive triangle-shaped autocorrelation function. Its cyclic extensions give even worse trapezoidal outputs. However, extending the sequence by $\mathbf{c}' = \{1, 1, 1, -3, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, -3, 1, 1, 1\}$ gives a \mathbf{c} -matched filter output of $\{\dots 3, 0, 0, 0, 0, 4, 0, 0, 0, 0, 3 \dots\}$. At the price of transmit energy efficiency (just as in ZAZ sequences), these extensions lead to improvements in the important sidelobes of the system.

4.2.1 Generalized extensions

The general case is typified in Figure 4.4. The right-extension \mathbf{u} , of any sequence \mathbf{c} giving perfect correlation, is given by the solution to the series of first-order linear equations

$$r_{cc}(n) + \sum_{i=0}^{n-1} u_i c_{L-n+i} = 0 \quad (4.10)$$

for $n = 1, 2, \dots$, where r_{cc} is the aperiodic autocorrelation of \mathbf{c} , and all sequences are zero-indexed. The z -transform of the sequence of solutions is

$$U(z) = -\frac{R_+(z)}{C^*(z)} \quad (4.11)$$

where $C^*(z)$ corresponds to the time-reversed version of \mathbf{c} , and $R_+(z)$ is the positive side of the autocorrelation function $r_{cc}(1) + r_{cc}(2)z^{-1} + r_{cc}(3)z^{-2} \dots$. In other words, \mathbf{u} is the output of an IIR filter with transfer function

$$H_u(z) = -\frac{1}{C^*(z)} \quad (4.12)$$

with $r_+(n)$ as input. We note that the infinite-length sequence \mathbf{u} thus generated can be truncated at any point without affecting the correlation zone built up to that point. Similarly, the left extension \mathbf{v} (which develops in reverse time) satisfies

$$V^*(z) = -\frac{R_+(z)}{C(z)}, \quad (4.13)$$

with a corresponding transfer function

$$H_v(z) = -\frac{1}{C(z)}. \quad (4.14)$$

Generally, unless all poles of H_v and H_u lie on the unit circle, either \mathbf{u} or \mathbf{v} , or both, will exponentially diverge to infinity. If one of $H_u(z)$ or $H_v(z)$ is stable (that is, all poles are within the unit circle), then the other is unstable, because reversing its denominator inverts its poles. If one of $H_u(z)$ or $H_v(z)$ contains poles both inside and outside the unit circle, then the other also holds that property and both filters are unstable.

Because the \mathbf{c} sequences in \mathbf{W} are native ZAZ sequences, the initial values of \mathbf{v}^* and \mathbf{u} must simply be the cyclic extensions of \mathbf{c} . At the limit of the zero autocorrelation zone, both \mathbf{v}^* and \mathbf{u} must differ from the cyclic extension. The generalized extensions of some sequences might diverge to infinity quickly and are thus impractical. However, we shall see that the \mathbf{W} sequences produce enough pole-zero cancelation effect to make interesting extensions which do not produce exorbitant transmit energies.

Two aspects of transmit waveform design are important in practical applications. The average transmit power, which is the product of the transmit waveform energy and the radar pulse repetition rate, determines the power generation or battery requirements for the transmitter platform. In some systems (such as, a lightweight, electric UAV transmitter), a doubling of transmitter power is quite significant because it will significantly the necessary size and weight of the battery for a given mission time span. The transmit waveform energy must be such that the signal to noise ratio (SNR) at the output of the matched-filter receiver is acceptable. The ZAC, and generalized ZAC, extensions do not materially affect receiver SNR; such extensions buy excellent autocorrelation performance for the price of additional transmit energy. Maximum instantaneous transmit power is also important for transmit-system dynamic-range circuit designs, covert activity considerations, and legal constraints.

A second design aspect of the transmit waveform design is its absolute length. The fundamental advantage of standard matched-filter pulse compression in radar is that it allows good receiver SNR and approximately delta-function-like resolution while spreading out the actual transmit energy in time [3]. One disadvantage of this standard scheme is that monostatic systems (or multistatic systems in close proximity) usually cannot usefully receive while transmitting; such radars are effectively blind to objects closer than about half of the length of the transmit waveform in

the air. When this blind spot is an important design consideration, then it is advantageous to keep the transmit sequence short, while (presumably) maintaining or improving other system design goals.

To assess these tradeoffs, it is convenient to define the transmit energy of the waveform so constructed to give zero correlation zone within a time offset of $\pm\tau$:

$$E(\tau) = \sum_{i=0}^{L-1} c_i^2 + \sum_{i=0}^{\tau-1} u_i^2 + v_i^2 \quad (4.15)$$

allowing a decibel figure of merit $dB_0(\tau) = 10 \log_{10}\{E(\tau)/E(0)\}$. An analysis of dB_0 for all 128 of the length-128 sequences of \mathbf{W}_7 is shown in Figure 4.5. The figure shows that all such sequences exhibit a 1.76 dB increase at the original ZAZ limit $\tau = L/4 = 32$. Subsequently, the energy increases at a higher rate, reaching 7.78 dB (for some sequences) at $\tau = 63$.

Figure 4.6 adds insight into the typical behavior of the best-case generalized-extension sequences with $\tau = L/2 - 1$. Up to $\tau = L/4$, the extension is cyclic and therefore repeats antipodal values of amplitude ± 1 . Between $\tau = L/4$ and $\tau = L/2$, the generalized extension exhibits antipodal values with amplitude ± 3 . Under these conditions, length- L sequences cyclically extended and truncated to $\tau = \pm L/4$ in the standard way exhibit zero autocorrelation up to $\tau = \pm L/4$ and burn a relative power of $dB_0(L/4) = 10 \log_{10} \frac{3}{2} \approx 1.8$ dB. The same sequences further extended to $\tau = L/2 - 1$ by generalized extension burn a relative power of $dB_0(L/2 - 1) = 10 \log_{10}(6 - 18/L) \approx 7.8$ dB. Therefore, the effective correlation zone of some \mathbf{W} sequences can be doubled for the price of about 6 dB in transmit power using generalized extensions.

The zone can also be doubled, in the standard ZAC style, by doubling the fundamental sequence length L , and the complexity of the receiver, at the price of 3 dB

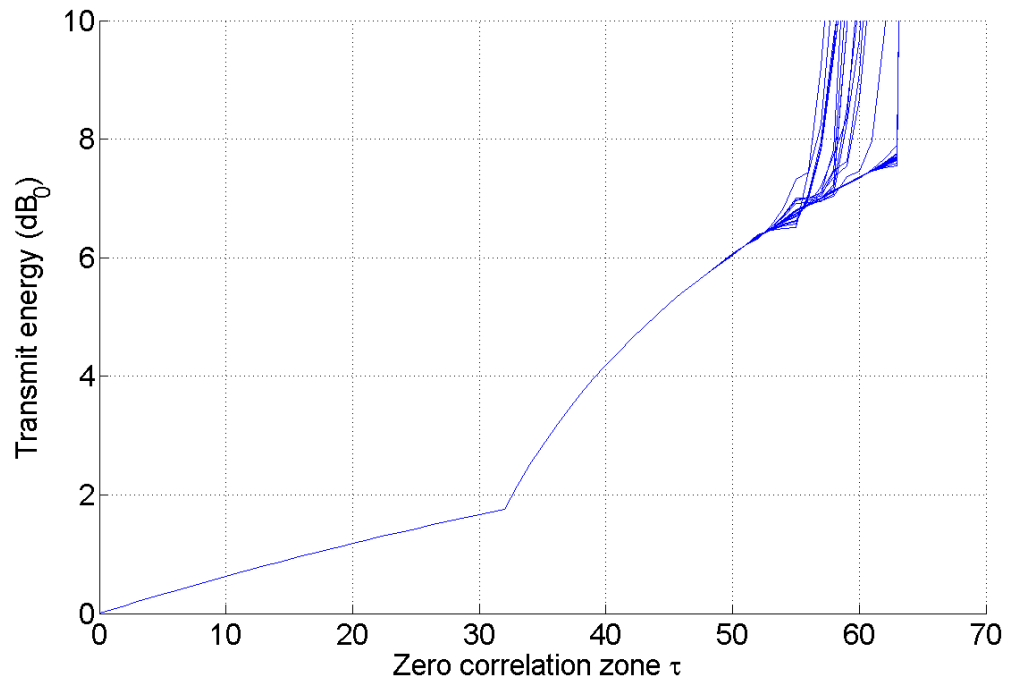


Figure 4.5: The transmit energy costs versus designed zero correlation zone τ for all sequences in \mathbf{W}_7 .

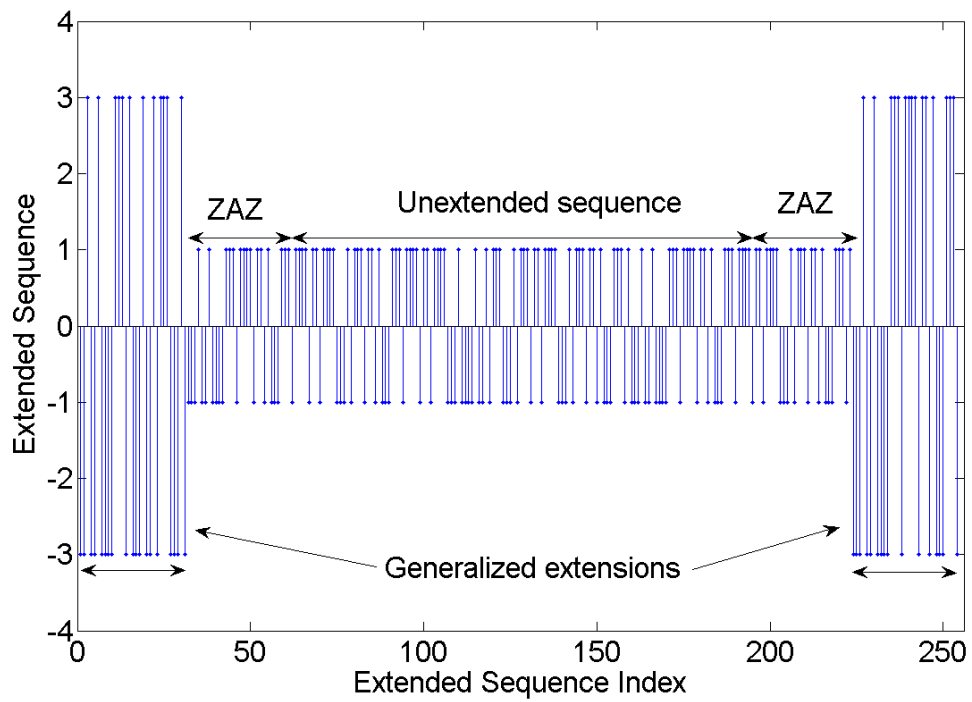


Figure 4.6: Generalized extensions for the first sequence of \mathbf{W}_7 up to $\tau = 63$.

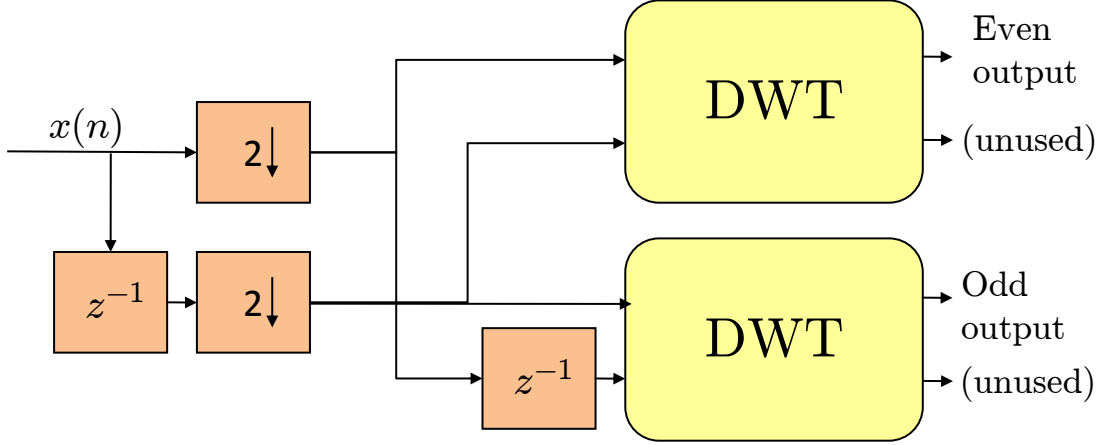


Figure 4.7: An efficient pipelined half-rate implementation of the matched filter of a sequence \mathbf{c} .

in transmit power. Therefore, the overall tradeoff due to the doubling of the zero correlation zone of the new sequences is the preservation of the receiver filtering complexity by doubling the transmit power. Overall, the new generalized-ZAC sequences are shorter given the same ZAC width, which is sometimes a benefit in monostatic radars, which usually cannot simultaneously transmit and receive information.

4.2.2 Implementation

The full receiver implementation for the sequences of this section is shown in Figure 4.7. All circuit elements downstream of the input deserializers operate at half of the input sample rate. The circuit uses one DWT to derive the even matched filter outputs, and another DWT to derive the odd matched filter outputs. The circuit requires $4 \log_2 L$ addition operations and $L - 1$ delay elements, representing an efficient pipelined implementation suitable for high-speed applications.

An interesting oddity in the circuit of Figure 4.7 is that half of the DWT outputs

are unused. It seems natural to ask whether there exists a transmit waveform for which a single DWT (using both half-rate outputs) would provide full-rate resolution for point reflectors.

4.3 Efficient use of the second output

In this section, we investigate the uses of a modified transmit sequence \mathbf{e} , which is derived from \mathbf{c} and \mathbf{d} in z -transform notation as

$$E(z) = C(z) + z^{-1}D(z). \quad (4.16)$$

The resulting sequence is ternary, $e_i \in \{0, \pm 2\}$, except for the outermost elements, which are $e_{0,L} \in \{\pm 1\}$.

The motivation for this construction is as follows. Assume for the moment that \mathbf{c} and \mathbf{d} are orthogonal under all shifts. Then, if \mathbf{e} is presented to the receiver with any even shift, then by construction, the output of the downsampled \mathbf{c} -matched filter will be a delta function, and by our temporary assumption of perfect orthogonality, the downsampled \mathbf{d} -matched filter output is zero. Conversely, if \mathbf{e} is presented to the receiver with an odd shift, then the roles of the \mathbf{c} and \mathbf{d} filters are reversed. Therefore, the \mathbf{e} construction has tricked the DWT decomposition into a shift invariance, in which the downsampled \mathbf{c} outputs service reflectors at even range offsets, and the downsampled \mathbf{d} outputs service reflectors at odd range offsets.

The advantage of this construction is that the half-rate dual output of a single DWT may be viewed as deserialized versions of a virtual full-rate correlation filter. An additional feature of interest is that the DWT outputs are simultaneously interpretable as correlation filter outputs, and also form a basis for the input waveforms.

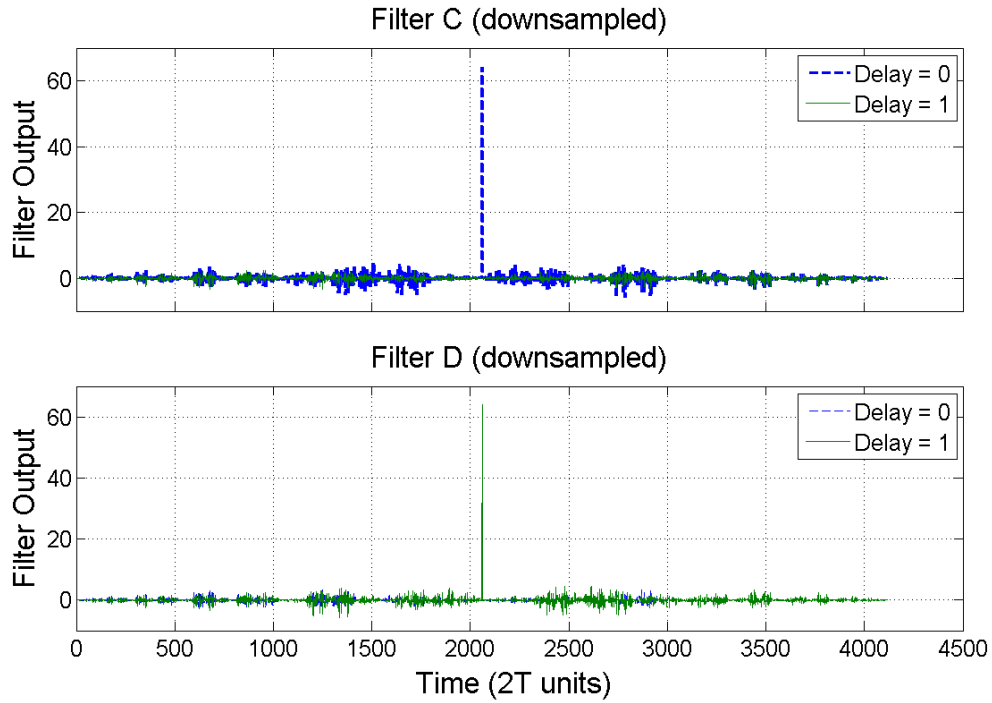


Figure 4.8: DWT decomposition of \mathbf{e} with $L = 2^{12}$ for different point reflector locations.

The simplest method to approximate orthogonality (under shifts) of \mathbf{c} and \mathbf{d} is to make the sequences long and to depend on the behavior of the aperiodic cross correlation between the sequences. Figure 4.8 shows the decomposition of \mathbf{e} using $L = 2^{12}$ for both even and odd delays. For even delay, the downsampled \mathbf{c} filter approximates a delta function, while the downsampled \mathbf{d} sees “noise” which is close to zero. For odd input delay, the \mathbf{d} filter approximates a delta function, while the \mathbf{c} sees “noise”. The maximum sidelobes in this case are approximately 25 dB down from the peak.

The maximum sidelobe seen over both filters and both delay conditions is the maximum value of the aperiodic cross correlation between \mathbf{c} and \mathbf{d} . Figure 4.9 shows the maximum sidelobe performance versus L for the best of the sequences of \mathbf{W} .

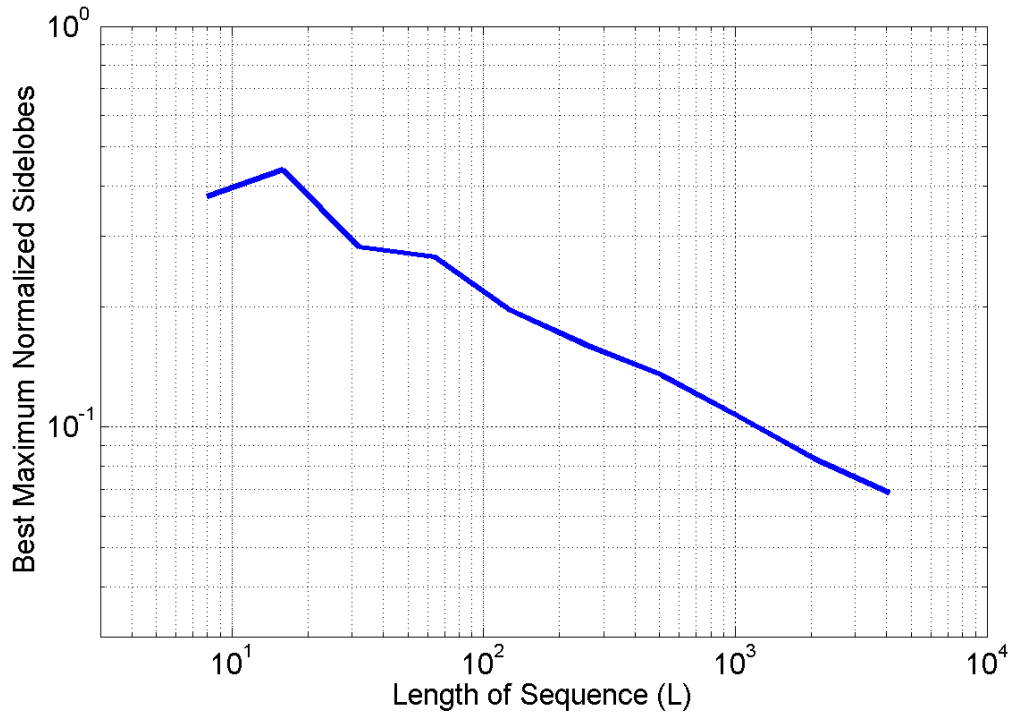


Figure 4.9: Normalized maximum sidelobes for DWT decomposition of \mathbf{e} versus sequence length.

Clearly, the orthogonality improves with sequence length, but perhaps only at a modest rate.

The energy in the transmit waveform is $\sum_i (c_{i+1} + d_i)^2 \approx 2L$ for large L , because the cross correlation $\sum_i c_{i+1}d_i$ has small magnitude (1 or -3) for these sequences. Therefore, the simplification in receive circuitry proposed in this section is once again bought for the price of 3 dB in transmit power. The implementation of the receive filters require a single DWT block in which both outputs are used symmetrically. The implementation complexity is $2 \log_2 L$ additions and $L/2 - 1$ delay elements.

4.4 Conclusions

We have shown that a DWT perspective on the ranging problem nicely summarizes previous research. Furthermore, the DWT implementation leads directly to an efficient pipelined implementation which is suitable for high-speed applications. Simple sequence constructions, using trivial modifications from the original research on complementary pairs and E -sequence characteristics (characteristics upon which discrete wavelets are based), were developed.

We have shown two different methods to trade off receiver complexity with transmit energy. In the first method, we developed a general method to extend the zero autocorrelation zone of any sequence. We found that this generalized construction was reasonable for the set of ZAZ sequences considered herein. In a second method, we exploited the approximate orthogonality under shifts of long E -sequences and their mates. This method addresses the problem of shift variance of the DWT for ranging applications. Both methods gain receive-complexity advantages at the cost of 3 dB in transmit energy. We suggest that the manifold tradeoff introduced by this work, between transmit power, sequence length, and receiver complexity might, best be used (for example) in a multistatic SAR system with a few high-power transmitters and many UAV receivers.

Chapter 5

Final remarks and future work

The UAV multistatic SAR problem, with its inherent bent to tradeoff simplicity of form and function in the UAV, for advanced signal processing at a central processing unit, is clearly a field producing many new problems, upon which future researchers and engineers might profitably graze. In this work, we have attempted to provide an approximate physical framework for simplified downlooking SAR systems, and then to apply some of the diverse mathematical tools now available to signal processors to illuminate and to simplify the radar tomography problem. While this work has achieved at least some of its goals, it has also exposed limitations in its physical and engineering assumptions.

It is appropriate, therefore, to recommend future extensions of this work, as well as new problems uncovered while developing this work.

1. Our simplifying assumption, that a reflectivity function $f(x, y)$ is wideband and isotropic, is approximate at best. Multistatic SAR seems well positioned to handle real-world reflective anisotropy because it is inherently capable of measuring a single piece of ground from many perspectives. Future work could develop a measurement framework, and generalized-image estimation methods,

in this area.

2. While simple UAVs might tend to move slowly, their necessarily-small antenna size means that modulation frequencies tend to be very high. Therefore, Doppler effects are expected to be significant in future UAV designs. Furthermore, in the downlooking systems developed here, the Doppler effect is not simple... some parts of the sampling ellipse have positive Doppler effect, and some parts have negative Doppler effect! Therefore, waveforms and methods designed to mitigate or exploit the generalized Doppler effects of this problem, are an important component of future research.
3. Problems of time synchronization in the proposed system are significant. We suggest variants upon the use of the primary (unreflected) waveform as a method to communicate timings between UAVs.
4. Although we used random sampling methods here, the full power of the compressed sensing discipline was not here applied to the multistatic SAR problem. The main results of this work apply to any image, and in particular, we did not take advantage of any sparse representation (or compressibility) of the image. Deeper investigation into the viability of physically-constrained projections to reduce the number of measurements is warranted.
5. We have seen that multistatic SAR problems can lead to very large problems in linear (and nonlinear) algebra. Further work on solution algorithms, possibly which take advantage of the sparseness of the \mathbf{A} matrix, is necessary for the success of any such system.

Bibliography

- [1] Robert O. Harger. *Synthetic Aperture Radar Systems: Theory and Design*. Electrical Science Series. Academic Press, 1970.
- [2] N.J. Redding. SAR image formation via inversion of Radon transforms. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 1, pages 13–16, 2004.
- [3] Merril I Skolnik. *Introduction to Radar Systems*. McGraw-Hill, 2002.
- [4] W. R. Madych. Radon's inversion formulas. *Transactions of the American Mathematical Society*, 356(11), January 2004.
- [5] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte Sächsische Akademie der Wissenschaften Leipzig*, 69:262–267, 1917.
- [6] Frank Natterer. Numerical methods in tomography. *Acta Numerica*, 1999.
- [7] P. Funk. Über Flächen mit lauter geschlossenen geodätischen Linien. *Mathematische Annalen*, 74:278–300, 1913.
- [8] Can Evren Yarman. *Inversion of Radon, exponential Radon, and Funk transforms based on harmonic analysis over groups*. PhD thesis, Rensselaer Polytechnic Institute, April 2006.
- [9] Can Evren Yarman and Birsen Yazici. Inversion of circular averages using the Funk transform. *ICASSP*, 2007.
- [10] Fritz John. *Plane Waves and Spherical Means Applied to Partial Differential Equations*. Number 2 in Interscience Tracts in Pure and Applied Mathematics. Interscience Publishers, 1955.
- [11] A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. *Journal of Applied Physics*, 34(9):2722–2727, 1963.
- [12] A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. ii. *Journal of Applied Physics*, 35(10):2908–2913, 1964.

- [13] A. M. Cormack. The Radon transform on a family of curves in the plane. ii. *Proceedings of the American Mathematical Society*, 86(2):293–298, Oct 1982.
- [14] A. M. Cormack. The Radon transform on a family of curves in the plane. *Proceedings of the American Mathematical Society*, 83(2):325–330, Oct 1981.
- [15] S. Helgason. *The Radon Transform*. Birkhauser, 1980.
- [16] Frank Natterer. *The Mathematics of Computerized Tomography*. John Wiley & Sons Ltd, 1986.
- [17] A E Yagle. Inversion of spherical means using geometric inversion and Radon transforms inverse problems. 8:949–964, 1992.
- [18] John A. Fawcett. Inversion of n-dimensional spherical averages. *SIAM Journal on Applied Mathematics*, 45(2):336–341, April 1985.
- [19] Lars-Erik Andersson. On the determination of a function from spherical averages. *SIAM Journal on Mathematical Analysis*, 19(1):214–232, 1988.
- [20] Lawrence Zalcman. Offbeat integral geometry. *The American Mathematical Monthly*, 87(3):161–175, March 1980.
- [21] H. Turner Laquer. The Pompeiu problem. *The American Mathematical Monthly*, 100(5):461–467, May 1993.
- [22] Leon Ehrenpreis. *The Universality of the Radon Transform*. Oxford Science Publications. Clarendon Press, 2003.
- [23] Robert S. Strichartz. Radon inversion – variations on a theme. *The American Mathematical Monthly*, 89(6):377–384, June/July 1982.
- [24] Robert S. Strichartz. Solutions of the problems in section 5 of “Radon inversion – variations on a theme”. *The American Mathematical Monthly*, 89(6):420–423, Jun/Jul 1982.
- [25] Ronald Bracewell. *Fourier Analysis and Imaging*. Kluwer, 2003.
- [26] Jack K. Cohen and Norman Bleistein. Velocity inversion procedure for acoustic waves. *Geophysics*, 44(6):1077–1087, June 1979.
- [27] Stephen J. Norton. Reconstruction of a two-dimensional reflecting medium over a circular domain: Exact solution. *Journal of the Acoustical Society of America*, 67(4):1266, April 1980.
- [28] Stephen J. Norton. Reconstruction of a reflectivity field from line integrals over circular paths. *Journal of the Acoustical Society of America*, 67(3):853, March 1980.

- [29] O'Brien J.D. Munson, D.C. and W.K. Jenkins. A tomographic formulation of spotlight-mode synthetic aperture radar. *Proceedings of the IEEE*, 71(8):917–925, 1983.
- [30] A. S. Milman. SAR imaging by ω -k migration. *International Journal of Remote Sensing*, 14, 1993.
- [31] Tummala S. Carrara, W. and R. Goodman. *Spotlight Synthetic Aperture Radar*. Artech House, 1995.
- [32] G. Franceschetti and R. Lanari. *Synthetic Aperture Radar Processing*. CRC Press, 1999.
- [33] Mehrdad Soumekh. *Synthetic Aperture Radar Signal Processing with Matlab Algorithms*. Wiley, 1999.
- [34] A. S. Milman. The hyperbolic geometry of SAR imaging. 2006.
- [35] N.J. Redding and T.M. Payne. Inverting the spherical Radon transform for 3D SAR image formation. In *Radar Conference, 2003. Proceedings of the International*, pages 466–471, 2003.
- [36] N. J. Redding and G. N. Newsam. Inverting the circular Radon transform. DTSSO Research Report DTSSO-RR-0211, August 2001.
- [37] Gestur Olafsson and Eric Todd Quinto, editors. *The Radon Transform, Inverse Problems, and Tomography*, volume 63 of *Proceedings of Symposia in Applied Mathematics*. American Mathematical Society, January 2006.
- [38] I. S. Gradshteyn and I. M. Ryzhik. *Tables of Integrals, Series, and Products*. Academic Press, 1965.
- [39] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions*. 1965.
- [40] J. J. Benedetto and A. I. Zayed, editors. *Sampling, Wavelets, and Tomography*. Birkhäuser, 2004.
- [41] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.
- [42] Yu Shkvarko and L. Leyva-Montiel. Theoretical aspects of Bayesian approach to aperture synthesis for radar imaging. In *Antennas and Propagation Society International Symposium, 2002. IEEE*, volume 4, pages 322–325, June 2002.
- [43] I.R. Joughin, D.P. Winebrenner, and D.B. Percival. Probability density functions for multilook polarimetric signatures. *Geoscience and Remote Sensing, IEEE Transactions on*, 32(3):562–574, May 1994.

- [44] I.R. Joughin, D.B. Percival, and D.P. Winebrenner. Maximum likelihood estimation of k distribution parameters for SAR data. *Geoscience and Remote Sensing, IEEE Transactions on*, 31(5):989–999, Sept. 1993.
- [45] A.L. Maffett and C.C. Wackerman. The modified beta density function as a model for synthetic aperture radar clutter statistics. *Geoscience and Remote Sensing, IEEE Transactions on*, 29(2):277–283, March 1991.
- [46] F. Chatalain and J.-Y. Tournet. Bivariate gamma distributions for multisensor SAR images. *ICASSP*, 2007.
- [47] R.S. Raghavan. A method for estimating parameters of K -distributed clutter. *Aerospace and Electronic Systems, IEEE Transactions on*, 27(2):238–246, March 1991.
- [48] K. Sigurd and J. P. How. UAV trajectory design using total field collision avoidance. *Proceedings of the AIAA Guidance, Navigation and Control Conference*, August 2004.
- [49] P.R. Chandler, M. Pachter, and S. Rasmussen. UAV cooperative control. *American Control Conference, 2001. Proceedings of the 2001*, 1:50–55 vol.1, 2001.
- [50] M. Rahim and S.M. Malaek. Intelligent operation using terrain following flight in unmanned aerial vehicles. *Aerospace Conference, 2007 IEEE*, pages 1–8, 3-10 March 2007.
- [51] G. Gu, P. R. Chandler, C. J. Schumacher, A. Sparks, and M. Pachter. Optimal cooperative sensing using a team of UAVs. *Aerospace and Electronic Systems, IEEE Transactions on*, 42(4):1446–1458, October 2006.
- [52] J. R. Higgins and R. L. Stens, editors. *Sampling Theory in Fourier Signal Analysis (Advanced Topics)*. Oxford University Press, 1999.
- [53] E. J. Candés and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies. Technical report, Applied and Computational Mathematics, Calif. Inst. Technol., 2004.
- [54] D. L. Donoho. Compressed sensing. *IEEE Trans. on Information Theory*, 52(4), 2006.
- [55] Pan Pan and D. Schonfeld. Image reconstruction and multidimensional field estimation from randomly scattered sensors. *Image Processing, IEEE Transactions on*, 17(1):94–99, Jan. 2008.
- [56] Jonathan D. Coker and Ahmed H. Tewfik. An elliptic formulation of omnidirectional multistatic SAR. *to appear*, 2008.

- [57] M. Knott. On the minimum efficiency of least squares. *Biometrika*, 62, 1975.
- [58] Zhaojun Bai and Gene H. Golub. Bounds for the trace of the inverse and the determinant of symmetric positive matrices. *Ann. Numer. Math.*, 4(1-4):29–38, 1997.
- [59] Moody T. Chu. Inverse eigenvalue problems. *SIAM Review*, 40(1):1–39, March 1998.
- [60] Moody T. Chu and Gene H. Golub. Structured inverse eigenvalue problems. *Acta Numerica*, pages 1–71, 2002.
- [61] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [62] L. R. Welch. Lower bounds on the maximum cross correlation of signals. *IEEE Transactions on Information Theory*, 20:398–399, 1974.
- [63] J.L. Massey. On Welch’s bound for the correlation of a sequence set. In *1991 IEEE International Symposium on Information Theory*, pages 385–385, 24–28 June 1991.
- [64] M. Stuart Lynn and William P. Timlake. Bounds for Perron eigenvectors and subdominant eigenvalues of positive matrices. *Linear Algebra and its Applications*, 2:143–152, 1969.
- [65] Jorma Kaarlo Merikoski. On a lower bound for the Perron eigenvalue. *BIT*, 19:39–42, 1979.
- [66] Jonathan D. Coker and Ahmed H. Tewfik. Multistatic SAR image reconstruction based on an elliptical-geometry Radon transform. *International Waveform Diversity and Design Conference*, 2007.
- [67] A. M. Tulino and S. Verdu. *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.
- [68] Taro Nagao and Toshiyuki Tanaka. Spectral density of sparse sample covariance matrices. *J. Phys. A: Math. Theor.*, 40, 2007.
- [69] Jonathan D. Coker and Ahmed H. Tewfik. Random sampling strategies in multistatic SAR. *ICASSP*, 2008.
- [70] Shaobing Scott Chen. *Basis Pursuit*. PhD thesis, Stanford University, November 1995.
- [71] E.J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, Feb. 2006.

- [72] ℓ_1 -MAGIC. www.l1-magic.org.
- [73] Robert Nowak Rui Castro, Jarvis Haupt and Gil Raz. Finding needles in noisy haystacks. *ICASSP*, 2008.
- [74] E. Candes and J. Romberg. Robust signal recovery from incomplete observations. *Image Processing, 2006 IEEE International Conference on*, pages 1281–1284, Oct. 2006.
- [75] Xinmin Deng and Pingzhi Fan. Spreading sequence sets with zero correlation zone. *Electronics Letters*, 36(11):993–994, 25 May 2000.
- [76] M. Nakamura H. Torii and N. Suehiro. A new class of zero-correlation zone sequences. *IEEE Transactions on Information Theory*, 50, 2004.
- [77] Daiyuan Peng and Pingzhi Fan. Bounds on aperiodic auto- and cross-correlations of binary sequences with low or zero correlation zone. In *Parallel and Distributed Computing, Applications and Technologies, 2003. PD-CAT'2003. Proceedings of the Fourth International Conference on*, pages 882–886, 27–29 Aug. 2003.
- [78] X.H. Tang and P.Z. Fan. A class of pseudonoise sequences over GF(P) with low correlation zone. *Information Theory, IEEE Transactions on*, 47(4):1644–1649, May 2001.
- [79] Xin Tong and Qiaoyan Wen. New constructions of ZCZ sequence set with large family size. *Signal Design and Its Applications in Communications, 2007. IWSDA 2007. 3rd International Workshop on*, pages 99–103, 23–27 Sept. 2007.
- [80] Fanxin Zeng and Lijia Ge. Some novel results on 1D and 2D sequences with zero correlation zone. In *Spread Spectrum Techniques and Applications, 2004 IEEE Eighth International Symposium on*, pages 934–938, 30 Aug.-2 Sept. 2004.
- [81] L.G. Weiss. Wavelets and wideband correlation processing. *Signal Processing Magazine, IEEE*, 11(1):13–32, 1994.
- [82] L. Rebollo-Neira and J. Fernandez-Rubio. On wideband deconvolution using wavelet transforms. *Signal Processing Letters, IEEE*, 4(7):207–209, 1997.
- [83] T. Le-Tien, H. Talhami, and D.T. Nguyen. Target signature extraction based on the continuous wavelet transform in ultra-wideband radar. *Electronics Letters*, 33(1):89–91, 1997.
- [84] J.-P. Ovarlez, L. Vignaud, J.-C. Castelli, M. Tria, and M. Benidir. Analysis of SAR images by multidimensional wavelet transform. *Radar, Sonar and Navigation, IEE Proceedings -*, 150(4):234–41, 2003.

- [85] Naiwei Wang, Yingguang Zhang, and Shunjun Wu. Radar waveform design and target detection using wavelets. In *Radar, 2001 CIE International Conference on, Proceedings*, pages 506–509, 2001.
- [86] Zhang Xiao-Ping, Tian Li-Sheng, and Peng Ying-Ning. Wavelet detectors for wideband radar signals. In *Radar, 1996. Proceedings., CIE International Conference of*, pages 289–292, 1996.
- [87] J.R. Sveinsson and J. Atli Benediktsson. Almost translation invariant wavelet transformations for speckle reduction of SAR images. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(10):2404–2408, 2003.
- [88] G. Welti. Quaternary codes for pulsed radar. *Information Theory, IEEE Transactions on*, 6(3):400–408, Jun 1960.
- [89] J. E. Golay. Complementary series. *IRE Transactions on Information Theory*, April 1961.
- [90] Y. Taki, H. Miyakawa, M. Hatori, and S. Namba. Even-shift orthogonal sequences. *Information Theory, IEEE Transactions on*, 15(2):295–300, Mar 1969.
- [91] M. Suehiro, N.; Hatori. N-shift cross-orthogonal sequences. *Information Theory, IEEE Transactions on*, 34(1):143–146, Jan 1988.
- [92] S. Z. Budisin. Efficient pulse compressor for Golay complementary sequences. *Electronics Letters*, 27, 1991.
- [93] B.M. Popovic. Efficient Golay correlator. *Electronics Letters*, 35, 1999.
- [94] M.C. Perez, J. Urena, A. Hernandez, A. Jimenez, W.P. Marnane, and F. Alvarez. Efficient real-time correlator for LS sequences. *Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on*, pages 1663–1668, 4-7 June 2007.
- [95] A. Rathinakumar and A.K. Chaturvedi. Mutually orthogonal sets of ZCZ sequences. *Electronics Letters*, 40(18):1133–1134, 2 Sept. 2004.
- [96] S. Stanczak, H. Boche, and M. Haardt. Are LAS-codes a miracle? *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, 1:589–593 vol.1, 2001.
- [97] Chin-Chong Tseng and C. Liu. Complementary sets of sequences. *Information Theory, IEEE Transactions on*, 18(5):644–652, Sep 1972.
- [98] Shu-Ming Tseng and M.R. Bell. Asynchronous multicarrier DS-CDMA using mutually orthogonal complementary sets of sequences. *Communications, IEEE Transactions on*, 48(1):53–59, Jan 2000.

- [99] Gilbert Strang. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, revised edition edition, 1997.
- [100] Ingrid Daubechies. Two-scale difference equations ii. *Siam J. Math. Anal.*, 23(4):1031, 1992.
- [101] Ingrid Daubechies. Two-scale difference equations i. *Siam J. Math. Anal.*, 22(5):1388, 1991.
- [102] P. P. Viadyanathan. *Multirate Systems and Filter Banks*. Prentice-Hall, 1993.
- [103] Jonathan D. Coker and Ahmed H. Tewfik. Characteristic phase E-sequences in efficient pulse compression methods using discrete wavelet decomposition. *ICASSP*, 2007.
- [104] R. Appuswamy and A.K. Chaturvedi. A new framework for constructing mutually orthogonal complementary sets and ZCZ sequences. *Information Theory, IEEE Transactions on*, 52(8):3817–3826, Aug. 2006.
- [105] T.C. Denk and K.K. Parhi. VLSI architectures for lattice structure based orthonormal discrete wavelet transforms. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on* [see also *Circuits and Systems II: Express Briefs, IEEE Transactions on*], 44(2):129–132, Feb 1997.
- [106] So Ryoung Park, Ickho Song, and Hyoungmoon Kwon. DS/CDMA signature sequences based on PR-QMF banks. *Signal Processing, IEEE Transactions on* [see also *Acoustics, Speech, and Signal Processing, IEEE Transactions on*], 50(12):3043–3054, Dec 2002.