DETECTION THEORY IN ULTRASONIC IMAGING

BY

ROGER JAMES ZEMP

B.Sc. (Physics, University of Alberta)

M.A.Sc. (Electrical-Computer Engineering, University of Toronto)

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Michael F. Insana

Craig K. Abbey

Jinyi Qi

Committee in Charge

2004

TABLE OF CONTENTS

LIST OF FIGURES AND PHOTOGRAPHS

LIST OF TABLES

1. INTRODUCTION

1.1. Cancer and Early Detection

1.1.1. Breast Cancer

Among women, breast cancer is one of the most commonly diagnosed types of cancer and is the second leading cause of cancer deaths after lung cancer. A National Cancer Institute (NCI) report estimates that about 1 in 8 women in the United States (approximately 13.3 percent) will develop breast cancer during her lifetime.[1] Much research has gone into understanding the basic disease mechanisms, as well as into ways to detect cancers early and treat them effectively. Our contribution will be to improve ultrasonic imaging technology so as to impact early detection.

Before going forward with the core material of our research it is appropriate to know some basic background on breast cancer. This can help us design strategies for imaging and early detection. The most common type of breast cancer is ductal carcinoma, which originates in milk ducts. Another type is lobular carcinoma, which originates in the lobules. We briefly discuss the current model for ductal carcinoma pathogenesis. Focal abnormalities in epithelial cells undergo malignant transformation due to carcinogens or other genetic or hormonal signaling. The apoptotic (programmed cell death) signalling pathways of malignant cells are altered so that cells grow in an uncontrolled way. The cellular microstructure is altered in a way that is recognizable by pathologists. The changes in microstructure also change ultrasonic echo signal properties - and may be a source of contrast. At first the malignant

cell population is contained within the milk ducts. At this stage the neoplasm is called a ductal carcinoma in situ. Signalling between the cancer cells and the surrounding stromal tissue can induce a desmoplastic effect (formation of a collagenous capsule thought to be responsible for tumor stiffness). Dangerously, Matrix Metaloproteinases (MMP's) and other signalling molecules generated by the neoplastic tissue can induce breakdown of the milk duct epithelium and lead to invasion of the surrounding tissue. At this point the cancer is called an invasive ductal carcinoma. This invasion is often accompanied by production of vascular endothelial growth factor and signaling molecules such as $VEGF$ and $\alpha_v\beta_3$ which initiate growth of new blood vessels (angiogenesis) to keep up with increased metabolic demand of the rampant cell growth. Also accompanying the disease progression is the decrease in the number of desmosomes and other focal adhesion molecules responsible for cellular adhesion and connectivity. The danger of invasion is thus that malignantly transformed cells can break away from the tumor into the bloodstream and migrate to remote parts of the body to seed new tumors.

1.1.2. Early Detection

Early detection of breast cancer can save lives. The American Cancer Society (ACS) report on the 5-year survival rates (percentage of patients who live at least 5 years after their cancer is diagnosed) for different cancer stage is summarized in Table 1. Stage 0 is ductal carcinoma in situ, and stages I through IV (1-4) correspond to increasing degrees of severity, including the tumor size and extent of metastasis.

Table 1. Breast Cancer Survival by Stage

Stage	5-year Survival Rate
0	100%
I	98%
IIA	88%
IIB	76%
IIIA	56%
IIIB	49%
IV	16%

The ACS also predicts an increase in survival rate from 82% to 95% if widespread screening is implemented as recommended. Detecting the breast cancer at an early stage, before cells have begun to metastasize, is necessary to increase survival rate and reduce cancer death.

Significant research has been performed to find methods which are effective at early detection. Methods that are both sensitive (can detect a high percent of cancers) and specific (have low false positive rates) are desired. The cost of low sensitivity could be a woman's life. The cost of low specificity are associated with the large-scale costs of large numbers of biopsies including risk of infection, financial costs to society and individuals, physical discomfort, and mental trauma of women who have a false diagnosis of cancer. Currently, the most common types of non-invasive methods for early detection include manual palpation and imaging.

Palpation is an essential diagnostic method for detecting neoplastic growth. In manual palpation, a hand presses onto the skin surface of a patient. Cancers are typically more stiff than the surrounding tissues, hence one can sometimes feel a 'lump' - which may (or may not) be a tumor. The problem is that some palpable lesions are already too big (typically 1 cm or larger). Ideally one would like to detect cancers when they are as small as possible to prevent growth and metastasis. Strain imaging techniques are meant to image the stiffness properties of tissues non-invasively in a way that is more sensitive than manual palpation, and these techniques are currently under investigation by a number of groups.[2–4]

Currently X-ray mammography is the gold standard imaging technique used for breast cancer screening. Mammograms are projection images through the breast that reveal contrast on the basis of the attenuation of properties of tissues, which are linked to density and atomic number. Suspicious masses and microcalcifications can be signs of cancer. Advantages of mammography include high specificity and sensitivity. Disadvantages of mammography include physical discomfort due to breast compression and the risks associated with the use of ionizing radiation. The role of other imaging techniques in the detection and management of malignant disease will be discussed shortly.

An understanding of risk factors and epidemiological studies is important because women who are at higher risk and their physicians may need hightened awareness to detect cancer as early as possible when treatment is most effective. Radiologists who know a women is at great risk for breast cancer may be more inclined to send

the patient for biopsy when viewing a suspicious region in an image. Risk factors for breast cancer include both genetic (e.g. BRCA1 or BRCA2 mutations, Li-Fraumeni syndrome due to TP53 mutations, and Cowden syndrome due to PTEN mutations[5]) and non-genetic factors including age, reproductive and menstrual history, hormone therapy, radiation exposure, mammographic breast density, lifestyle factors, and history of breast disease.

1.1.3. Diagnostic B-Mode Ultrasound

A controversial Danish study used radiation dose calculations and mortality data from the Hiroshima and Nagasaki tragedies to predict the risk versus benefit of mammography as an early detection modality.[6] They concluded that the radiation doses from mammograms could be causing as many cancers as are being detected successfully with early screening (within statistical errors).

It would be a significant contribution to the medical community to develop an early detection system that could equal or exceed the detection performance of mammography systems using non-ionizing radiation. Ultrasound systems, with some improvements, could satisfy this requirement.

Ultrasound systems use safer high frequency sound waves rather than ionizing radiation. Additionally ultrasound systems are portable, relatively low cost, and offer realtime imaging capabilities that are invaluable in the clinical environment.

Ultrasound systems are already playing a crucial role in the early detection, diagnosis, and management of malignant disease. Advancements in ultrasound technology

have expanded the uses for ultrasonography in the evaluation of the breast. Breast masses can be delineated by specific ultrasonographic characteristics that allow them to be categorized according to their relative risk of being malignant. When combined with physical examination and mammography, breast ultrasonography can decrease additional radiation exposure associated with repeat mammograms. It also can lower the cost of evaluation of the breast, and it often reduces the number of open biopsies.[7] In addition, ultrasound can be used for definitive pathologic diagnosis by guiding fine-needle aspiration and core biopsy, as well as facilitating preoperative needle localization for excisional biopsy.

Sonographers search for signs of malignancy when a lesion is discovered. Fluid filled cysts are benign and can often be distinguished from malignancies on the basis of echogenicity (cysts have very few scatterers inside), acoustic enhancement at the proximal surface of the cyst borders - (partially due to an acoustic lens effects caused by the spherical shape of the cyst boundary). Fibroadenomas are benign growths with complex shapes that sometimes resemble tumors in sonograms. Tumors are almost always hypo-echoic, appearing as dark regions in a bright backgrounds. Additionally, malignant tumors often have spiculated boundaries, and are oblong the growth pattern following the natural tract of milk ducts. Malignant tumors have typically ruptured their stromal capsule and the tumor cells have invaded surrounding tissues - thus malignant tumors (e.g. IDC) often have a more diffuse boundary than ductal carcinoma in situ.

In one study by Chen et. al.[8] they evaluated the diagnostic ability of ultrasound

and defined the sonographic features of symptomatic intraductal and invasive breast carcinoma. They concluded that differentiation of ultrasonic features of intraductal and invasive carcinoma can be based on the internal hypo-echogenicity, loss of bilateral edge shadowing, posterior acoustic transmission, irregular shape and non-uniform internal echo texture. The study, however was retrospective, and further investigation is needed to examine the role of using these features for detection.

Irwig *et. al.*[9] reviewed imaging technologies for screening applications and put ultrasound performance in context with mammography (including Full-Field Digital Mammography and Computer Aided Detection) and MRI. Kolb *et. al.*[10] conclude that ultrasound as an adjunct to mammography detects additional cancers, especially in women with radiographically dense breast tissue. Irwig *et. al.* point out however that sonography may introduce additional false positives. It is important to note, however, that these results are a combination of detection (i.e. is there something abnormal there?) and classification tasks (i.e. is the abnormality a cyst, fibroadenoma, ductal carcinoma in situ or an invasive ductal carcinoma?), and that their study is very technology dependent. Our point is that improved ultrasonic technology could result in improved diagnosis.

1.2. Image Quality

Good image quality is the design objective of every modality in medical imaging. Image quality affects the ability of physicians to diagnose and manage disease and determine a patient's state of health. Modalities on the frontier of medical imaging are

also helping us understand the mechanisms of disease progression - thus contributing to basic science research.

Traditionally image quality metrics such as contrast, contrast to noise ratio, spatial resolution, and echo signal-to-noise ratio have been used to characterize imaging system performance. There are however, almost always tradeoffs between these parameters, and the optimum balance among parameters is application dependent.

1.2.1. Task Performance

We follow a school of thought pioneered by many before us[11] that image quality should be defined in terms of the task at hand, and that task performance should and can be measured quantitatively.

For the medical diagnostic task of early detection of breast cancer, we may actually need to consider several sub-tasks: (1) Is there an abnormal growth present or isn't there? (2) If there is a growth present, is it a mass or a fluid filled cyst? (3) If it is a mass, is it a fibroadenoma (benign) or a carcinoma (malignant)? (4) If it is a carcinoma - what stage is it at - Ductal Carcinoma In Situ, or Invasive Ductal Carcinoma? (5) If it is an Invasive Ductal Carcinoma, what is the extent of metastasis?

Many of these sub-tasks are binary classification tasks - a clinician must decide between two categories of disease states based on the image data. In this dissertation, much of our focus will be on the most basic of binary classification tasks, detection tasks, where one must decide if a lesion is present or absent. It will turn out that the

formulation for analyzing many other binary classification tasks will be very similar to the detection case. One reason that we focus on detection problems is that this is the task most analyzed in other modalities. Thus there is a natural connection to a broader framework for analyzing and designing imaging systems that we will explore in the next chapters.

1.2.2. Observers and Decision Theory

In clinical situations, it is radiologists that make diagnoses on the basis of medical images and patient history. Computer aided diagnosis in mammography is becoming more common to assist in the detection process. We can further abstract the concept of a decision maker to a mathematical framework or computational algorithm that forms a scalar test statistic $t(\mathbf{g})$ from the image vector $\mathbf{g}$. This *observer* compares the test statistic with a threshold value t_o to make a decision. For example, if $t(\mathbf{g}) \geq t_0$ it may be decided that there is a lesion present, and if $t(\mathbf{g}) < t_o$ the opposite is decided. Over an ensemble of cases, a histogram of test statistic values for each hypothesis is formed. These histograms are governed by probability distributions, which in some cases can be modeled by analytical functions. When probability distributions for each hypothesis overlap, decision errors occur. The performance of an observer is analyzed by the probabilities of decision errors. A receiver operating characteristic (ROC) curve[12] is the gold standard for doing this. An ROC curve plots the true positive fractional (TPF) probability for all possible false positive (FPF) values. Curves with large area (close to 1) identify high performance diagnostic tests.

The area under the ROC curve is a common summary measure of observer performance. Another common measure is the observer signal to noise ratio, defined as the mean separation in the test statistic likelihoods normalized by the common standard deviation:

$$SNR_t = \frac{\overline{t}_1 - \overline{t}_0}{\sqrt{0.5(\sigma_1^2 + \sigma_0^2)}} \tag{1}$$

where $\overline{t}_1$, $\overline{t}_0$, σ_1^2, and σ_0^2 are the means and variances of the test statistic conditioned on the signal present and signal absent hypotheses, respectively.

ROC curves for diagnostic detection tasks are observer and imaging system dependent for a given population of patient data.

Relating to diagnostic image quality, there are several classes of mathematical or computational observers. One class is meant to mimic human observer performance. Human observer studies typically require radiologists or others to view large ensembles of images and make decisions on each. Often these images are simulated so that truth about a lesion present or absent is known - yet still these studies are time-consuming and expensive. A computational observer that mimics the human observer could greatly facilitate imaging research. For example, one may wish to test whether a processing algorithm conveys more diagnostic information to the human observer. Running a computational observer algorithm could be very valuable to assess performance for a given diagnostic task. Modeling human observer performance is a complex challenge - and one that we do not discuss very much in this dissertation - although we are concerned with measuring human observer performance.

Other classes of mathematical and computational observers are not meant to model human observer performance, but are rather concerned with ideal decision makers that are perhaps better than human observers. These observers are meant to extract the maximum possible task-dependent information from the data, which humans cannot.

1.3. The Bayesian Ideal Observer

The observer with performance given by the best of all possible ROC curves for a given system design is called the Bayesian ideal observer.[13] It has the highest TPF's for all given FPF's compared to all other observers. The Bayesian ideal observer is able to perform at the highest possible level because it has full knowledge of the probability laws governing the data. It uses as its test statistic the log-likelihood ratio, given as

$$\lambda(\mathbf{g}) = ln\left(\frac{pdf(\mathbf{g}|H_1)}{pdf(\mathbf{g}|H_0)}\right) \tag{2}$$

where H_1 and H_0 represent the hypothesis for the signal present and signal absent cases, respectively, and the pdfs represent probability density or mass functions. Monotonic transformations of the log-likelihood ratio can be shown to give identical performance to the log-likelihood.[14]

With prior probabilities and a model for decision error costs, a Bayesian decision maker can additionally choose the test statistic threshold values that can optimize the cost-versus-benefit tradeoffs for a given task. In this dissertation, we do not concern

ourselves with the threshold values, as we are more interested in the ideal observer for arbitrary threshold values.

The ideal observer is very important for optimizing medical imaging systems for diagnostic tasks. One reason for this is because the performance of the ideal observer reflects the upper bound on performance. Because the ideal observer is optimal, it uses all possible information in the raw data to make its decision. Thus the ideal observer processes the data in the optimal way to make its decision. Much can potentially be learned by mimicking the ideal observer's strategy. Another reason is that the ideal observer is a measure of diagnostic information from the detection system. Knowing that the ideal observer can process the data in the optimal way means that it can only perform to the extent that information is in the data. By studying how to optimize the performance of the ideal observer, we thus are studying how to improve diagnostic information content in the raw data. Importantly, this directly impacts system hardware design independent of the processing algorithms used to manipulate and display the data. The role of processing algorithms for diagnostic imaging tasks is to provide access to human observers to as much of the diagnostic information in the data as possible. By measuring the human observer performance (dependent on the system design, the processing algorithms, the display, and the observer) relative to the ideal observer, one can then assess the performance of the the processing and display methods. If the ideal observer has much better performance than human observers - this suggests an improved role for image processing.

1.4. Dissertation Objective

This dissertation is concerned with modeling the ideal observer for detection and discrimination tasks in ultrasonic imaging. In effect, a thorough understanding of the ideal observer could help us push the diagnostic limits of ultrasound to unprecedented levels - thus significantly aiding in the detection, diagnosis, staging, and management of malignant disease in the breast and in other pathologies.

Our approach is to model essential features of ultrasonic detection and discrimination tasks, including statistical object and noise models and realistic system models integrating a full physics-based approach. Equipped with such a model, we are poised to analytically and computationally study the form of the ideal observer test statistic for ultrasonic detection tasks. We then use analytical and Monte Carlo simulation approaches to assess the performance of system designs for detection tasks. The ideal observer analysis suggests instrumentation design strategies and has also led to some curious results concerning unfocused beams. From this analysis a new imaging technique emerged and was implemented on a commercial ultrasonic scanner. Thus we ultimately test some of our ideas experimentally. We show that these methods can significantly improve diagnostic image quality over dynamic receive processing methods currently used on most systems. Ultimately we hope that our efforts may contribute to improved early detection of breast cancer and impact other areas of clinical ultrasonography.

2. BACKGROUND AND LITERATURE REVIEW

2.1. Ultrasound Fundamentals

Later in the dissertation we will refer to a number of common techniques used in ultrasonic system designs. For this reason we give a brief outline of some fundamental principles of operation and techniques currently used in modern ultrasonic scanners.

2.1.1. B-mode Imaging

Ultrasonic imaging systems operate by sending high frequency sound waves into the body and receiving the scattered echo signals. Variations in density and compressibility of tissue micro- and macro-structure are responsible for the scattering and thus the variance of these quantities are responsible for the contrast observed in ultrasonic images.

Modern ultrasonic scanners use array transducers having a large number of small elements with widths typically on the order of the wavelength of the transmitted acoustic energy. A subset of elements are excited with programmable delays to simulate a geometrically focused lens. The sub-aperture transmits a pulse and the returning echoes are collected by the array elements and coherently summed focus the beam on receive. The focusing properties of the beam determine the lateral resolution of the imaging system. Axial resolution is determined by the pulse length - or more accurately the reciprocal of the pulse bandwidth.[?]

An A-scan line is the beamformed echo received along a single line of sight. Subapertures of elements can be electronically walked across the array to acquire multiple

A-scan lines, which are spatially arranged in a given scan geometry to form a B-mode image. B-mode images are typically envelope detected RF signals that are spatially registered, sometimes with additional image processing such as log-compression or histogram equalization[15] to increase the visual dynamic range.

Array fabrication and system interfacing for high frequency systems has proven to be more difficult, and sometimes single element transducers are mechanically scanned in place of electronic beam steering.[16]

Ultrasound images have a speckle texture that is characteristic of the tissue, the transmitted energy spectrum, and the acoustic focusing properties of the transducer.[17]

2.1.2. Dynamic Receive Focusing

For a single acoustic transmission, one chooses a single transmit focal location. However, on reception, one can adjust element delays dynamically to focus at each receive depth. In this way modern ultrasound scanners can significantly improve the depth of field, meaning that fine spatial resolution can be maintained throughout a large axial depth centered at the transmit focal length. This technique is called Dynamic Receive Focusing. The delay function is given in Eq. (8) of Ref..[18] See also Ref.[19] and it's references.

2.1.3. Aperture Growth

Additionally, to make the spatial resolution more uniform throughout the region of interest, systems often dynamically increase the size of the receive aperture with

increasing depth (or decrease the aperture with decreasing depth) to maintain an approximately constant f-number (defined as the ratio of the focal depth to the aperture size). This technique is called aperture growth. With aperture growth, the impulse response of the dynamically focused receive beam is approximately shift-invariant. Thus the net pulse-echo impulse response (IR) function is given by the 1-D spatial convolution of the IR of the received beam wiith the shift-varying IR of the transmit beam along the beam axis.

2.1.4. Apodization

Phase delays are applied to the echo signals on individual array elements to focus the beam. However, amplitude weighting of elements is also possible on both transmission and reception. Typical lateral amplitude profiles applied to apertures include Gaussian and Hanning window functions. Apodization reduces sidelobes generated from the discrete element approximation to a continuous transducer surface while widening the mainlobe; thus it increases contrast resolution at the cost of lower lateral spatial resolution.[20]

2.1.5. Synthetic Aperture Techniques

Ultrasonic imaging is in many ways similar to radar or sonar. A technique originating from radar called synthetic aperture imaging[21] can also be applied to ultrasonic imaging systems with similar advantages and disadvantages. Because lateral spatial resolution is proportional to the f-number, high flying aircraft or satellites need a very large aperture to attain good spatial resolution. Instead, they synthesize large

apertures by tranmitting pulses and receiving echoes along their flight path and later coherently combining echoes to reconstruct images of spatial resolution much greater than that possible with the physical aperture. In modern ultrasound systems, synthetic aperture techniques are not routinely used, in part because the computational burden required to process the data has been prohibitive for realtime applications, and in part because standard synthetic receive aperture techniques require transmission of small (point) sources - thus significantly reducing the energy per transmission compared to using a much larger physical aperture. If transmission and reception processes can take place on different spatially separated elements, and N transmit firings each have N receive echo signals (called a complete set), additional focusing can be done on transmission as well as upon reception.[22]

2.1.6. Coded Excitation

Traditionally pulses used for ultrasound imaging have been similar to Gaussian-modulated sinusoids. More recently, some systems have borrowed a technique from radar called coded excitation, which sends frequency coded waveforms such as chirp signals, and digital phase codes such as Barker codes, Golay codes, and Optimal codes.[18] These waveforms are meant to be much longer than pulses only a few cycles long but have comparable bandwidth, which means they have the potential to provide comparable spatial resolution and larger echo SNR. To recover the spatial resolution, the echoes must be decoded. Incompletely decoded echoes have large range lobes. These reduce contrast resolution just as beam sidelobes do. Codes enhance the echo

SNR in ultrasound images, and thus increase penetration depth in attenuating media, particularly at high frequencies.

2.1.7. Aims in this Dissertation

In the end sections of this dissertation we introduce a technique that improves on the spatial resolution and echo signal-to-noise of dynamic receive focusing, and is a generalization of coded excitation and synthetic aperture techniques. Additionally, the technique which we call *unfocused imaging* or *spatial coding* can be implemented in realtime on commercial scanners, with greater echo SNR compared to traditional synthetic receive aperture techniques. Unfocused imaging was a direct outcome of our research on the ideal observer for ultrasonic imaging in that it implements the strategies of the ideal observer for detection.

2.2. Image Quality Assessment in Other Modalities

Before embarking on the ideal observer for ultrasonic detection tasks, we note that much has been done for other imaging modalities. It is important to bring the to readers attention these contributions because in some cases it has significantly impacted the way engineers and researchers design and evaluate medical imaging systems. For example, photon imaging modalities use a concept called noise equivalent quanta (NEQ) that is directly related to the ideal observer for detection tasks in those modalities. NEQ, a reflection of the fraction of photons that carry information per unit spatial frequency, has become a widely used figure of merit for assessing the performance of these modalities. Smith and Wagner showed that speckle spots are

the ultrasonic equivalent of information carrying photons in x-ray or optical techniques under very limited conditions, such that the NEQ concept can also be applied to ultrasound. Our research has sought to generalize the NEQ concept to a broader range of imaging conditions.

To begin, we first review the common language of image quality assessment among many modalities, and introduce widely used figures of merit. We then proceed by connecting these metrics to signal detection theory.

2.2.1. Linear Systems Models

A unifying approach to modeling many imaging modalities is to model the image formation process as a linear system,

$$\mathbf{g} = \mathcal{H}f + \mathbf{n} \tag{3}$$

where $\mathbf{g}$ is the vector of image data, f represents the object function defined over space, and $\mathbf{n}$ is additive noise. $\mathcal{H}$ is a continuous-to-discrete operator describing how the system maps object functions into data samples. Often it will be useful and reasonable to assume that the object can be approximated as discrete (represented by a vector $\mathbf{f}$) and that $\mathcal{H}$ can be approximated as a matrix $\mathbf{H}$. Regardless of the dimensionality of the object or image data, the elements may be concatonated into a column vector in a rastorized fashion.

In a subsequent chapter we rigorously elucidate the precise physical model underlying the ultrasonic version of this model.

The model is important because it lends itself to a great deal of powerful analysis tools, matrix algebra, and stochastic analysis that will be essential to understand and push the performance limits of imaging systems.

2.2.2. Linear Shift Invariance

Over small regions, called isoplanatic patches, it may be assumed that the imaging system is linear shift invariant so that a convolution model may be assumed. In this case, the matrix **H** may be approximated as a block-circulant (rather than just a block-Toeplitz) matrix. Alternatively, the linear system can be written in continuous form as

$$g(\mathbf{x}) = h(\mathbf{x}) * f(\mathbf{x}) + n(\mathbf{x}) \tag{4}$$

where now the arguments of the relevant quantities are now spatial vectors (related linearly to the time vectors).

2.2.3. Modulation Transfer Function

An important quantity used in the medical imaging literature is the Modulation Transfer Function (MTF), defined as[23]

$$MTF(\mathbf{u}) = \frac{|H(\mathbf{u})|}{|H(0)|} \tag{5}$$

where $H(\mathbf{u})$ is the spatial Fourier Transform of the linear shift-invariant impulse response function $h(\mathbf{x})$, and $\mathbf{u}$ is a spatial frequency vector. The maximum value of H in phase insensitive photon imaging modalities is usually at DC. The DC value, $|H(0)|$, is sometimes referred to as the large area contrast transfer (sometimes written

as G), since for large targets, it reflects the ratio of the image contrast (C_{out}) to the object contrast (C_{in}), i.e. $|H(0)| = C_{out}/C_{in}$. [Note that object contrast C_{in} and image contrast C_{out} can have different units].

2.2.4. Noise Equivalent Quanta for Photon Imaging Modalities

A widely used figure of merit for designing photon imaging systems is Noise Equivalent Quanta, or NEQ, originally proposed by Shaw[24] which is a spatial frequency measure of the fraction of photons that carry information. "It gives the number of photons that would produce the same SNR given by an ideal detector on an absolute scale". It can be written as

$$NEQ(\mathbf{u}) = \frac{|\overline{q}H(\mathbf{u})|^2}{NPS(\mathbf{u})} \tag{6}$$

where $\overline{q}$ is the photon fluence, defined as the average number of photons per unit area incident on the detector surface, and $NPS(\mathbf{u}) = \langle |N(\mathbf{u})|^2 \rangle$ is the noise power spectrum of the additive noise process n, which we will sometimes write as $S_n(\mathbf{u})$. For a perfect system $NEQ = \overline{q}$ since the MTF is 1 over the entire spatial frequency range, and the noise power spectrum has variance equal to the mean fluence (Poisson Random Process). NEQ can be seen as a summary measure of potential system performance (as opposed to task performance) when you consider it is comprised of factors related to image contrast, $\overline{q}|H(0)|^2$, spatial resolution, $|H(\mathbf{u})|^2/|H(0)|^2$, and noise power spectrum $NPS(u)$.

2.2.5. Detective Quantum Efficiency

Detective quantum efficiency (DQE) is essentially a normalized version of NEQ such that a perfect system would give a DQE of one,

$$DQE(\mathbf{u}) = \frac{NEQ(\mathbf{u})}{\overline{q}}, \tag{7}$$

where $0 \leq DQE \leq 1$.

2.3. Signal Detection Models in Other Modalities

2.3.1. Signal Known Exactly and Signal Known Statistically Tasks

Task performance assessment includes system performance only in the context of a task. Often we simplify the task in order to express the response of a model observer in closed form. The most basic form of signal detection task is where the exact signal to be detected is known *a priori*. For example, in medical diagnosis, we know the lesion location, shape, and the precise signal; our job is only to determine if the signal is present. By this we mean that we know the lesion location, shape, and the precise signal from the lesion. We denote such situations with the label SKE (Signal Known Exactly) as opposed to Signal Known Statistically (SKS).[23] Ultrasound detection tasks that we consider in this thesis have known location but may have other object properties characterized statistically.

2.3.2. Background Known Exactly and Background Known Statistically Tasks

Object signals $\mathbf{f}_s$ in detection theory are often assumed to be additive over some background $\mathbf{f}_b$ so that the total object is $\mathbf{f} = \mathbf{f}_s + \mathbf{f}_b$. The background process may be deterministic and known exactly, in which case it is referred to as a Background Known Exactly (BKE) task or it may be characterized stochastically, and referred to as Background Known Statistically (BKS) tasks.

2.3.3. The Rose Model

Consider a SKE, BKE detection task in a photon imaging modality. Assume that there are n_s photons per unit area incident on the detector during the exposure time, and that the target is a disk of area A. The total mean number of photons in the target area is N_s and the difference relative to the background is $\langle \Delta N_s \rangle = A\langle \Delta n_s \rangle$. Here $\langle \cdot \rangle$ is the ensemble mean of the argument. Rose defined contrast C as early as the 1940s as[25]

$$C = \frac{\langle \Delta N_s \rangle}{\langle N_b \rangle} = \frac{\langle \Delta n_s \rangle}{\langle n_b \rangle} \tag{8}$$

and detection signal-to-noise as the mean signal over the mean noise:

$$SNR_{Rose} = \frac{\langle \Delta N_s \rangle}{\sqrt{\langle N_b \rangle}} = \frac{A\langle \Delta n_s \rangle}{\sqrt{A\langle n_b \rangle}} = C\sqrt{A\langle n_b \rangle}. \tag{9}$$

Like NEQ, SNR_{Rose} is a system performance metric; it depends on image contrast C, spatial resolution (in this case given by the target size $\sqrt{(}A)$ since an ideal impulse response is assumed), and a noise $\langle n_b \rangle = (\sigma_n^2/\langle n \rangle)^{-1}$. Although the Rose model

excludes many important properties, it was the first attempt to quantify the concept of image quality.

When the number of photons is large, the underlying Poisson process can be thought of as Gaussian, and the signal-to-noise SNR_{Rose} can quantitatively parameterize the ROC detection curve[12] for the test statistic t defined as the integrated number of photons over the lesion area, and is equivalent to the so-called d' value, defined by[26,14]

$$d' = \frac{t_s - t_b}{\sqrt{0.5(\sigma_{t_s}^2 + \sigma_{t_b}^2)}} \tag{10}$$

A low contrast assumption is needed to make the formal connection between SNR_{Rose} and d'. The Rose model was important because it connected contrast, area, and photon rates together for the task of detection of disk targets.

2.4. Ideal Observers Models In Other Modalities

2.4.1. Wagner-Brown Model

Wagner and Brown[27] extended the Rose model to include the effects of the system spatial resolution on the detection task. In particular their analysis stems from an analysis of the SKE, BKE ideal observer for Gaussian signals. For the first time, they connected a specific detection task with previous notions of MTF, NEQ, contrast, etc. into a single expression for the ideal observer SNR. Their theoretic framework has been applied to most major medical imaging modalities[27] - with one noticeable exception being ultrasound systems - which has been addressed partially in other work,[28,29] and is the topic of this dissertation.

For SKE, BKE detection tasks given Gaussian distributed signals, the prewhitening matched filter

$$\lambda = \Delta\overline{\mathbf{g}}^t \mathbf{K}_n^{-1} \mathbf{g} \tag{11}$$

is the test statistic of the ideal observer, where $\mathbf{g}$ is the data vector, and $\mathbf{K}_n = \langle \mathbf{n}\mathbf{n}^t \rangle$ is the covariance vector of the zero-mean noise process, which we typically assume is Gaussian white ($\mathbf{K}_n = \sigma_n^2 \mathbf{I}$) or colored noise. $\Delta\overline{\mathbf{g}} = \mathbf{H}\Delta\overline{\mathbf{f}}$ represents the difference signal between the target and the background, such that $\Delta\mathbf{f} = \mathbf{f}_s - \mathbf{f}_b$.

In this dissertation we give great emphasis to the strategy of the ideal observer as revealed by the form of its test statistic. The strategy of the ideal observer for this task is to first pre-whiten the data vector $\mathbf{g}$ by "un-doing" the noise correlations with the operation $\mathbf{K}_n^{-1}$ then to match filter by first multiplying with $\mathbf{H}^t$, then by integrating over the lesion template $\Delta\overline{\mathbf{f}}$.

The signal-to-noise ratio (performance) of this test statistic is given as

$$SNR^2_{Wagner-Brown} = \Delta\overline{\mathbf{g}}^t \mathbf{K}_n^{-1} \Delta\overline{\mathbf{g}} \tag{12}$$

If we assume linear shift-invariance, $\mathbf{H}$ can be thought of as block-circulant, and thus can be diagonalized by a Discrete Fourier Transform, written in matrix form as $\mathbf{F}$. Consequently, it can be written in continuous form as

$$SNR^2_{Wagner-Brown} = \int d\mathbf{u} \frac{|\Delta F(\mathbf{u})|^2 |H(\mathbf{u})|^2}{NPS(\mathbf{u})} \tag{13}$$

where $\Delta F(\mathbf{u})$ is the Fourier Transform of $\Delta f(\mathbf{x})$.

Importantly, this can also be written as

$$SNR^2_{Wagner-Brown} = \int d\mathbf{u} |\Delta F(\mathbf{u})|^2 NEQ(\mathbf{u}) \tag{14}$$

This equation decouples contributions from the system (NEQ) and the object (sometimes called the task $|\Delta F(\mathbf{u})|^2$). It connects formal task performance of the ideal observer with the widely used Noise Equivalent Quanta for characterizing imaging system performance for detection tasks - thus NEQ has a truly information-theoretic interpretation. We construct a similar unifying equation for ultrasonic imaging in later chapters. For an ideal system, the $SNR_{Wagner-Brown}$ reduces to SNR_{Rose}. When the system is not ideal, the framework adds additional information to the detection task about the system spatial resolution.

2.4.2. Other Key Literature

Harrison Barrett and his group have published an extensive number of papers on Objective Assessment of Image Quality and Ideal Observer Performance. Three key papers,[11] [30] and[13] in particular give a foundation for much of what we do in this dissertation. These papers give a rigorous treatment of the Non-prewhitening, Hotelling, and Ideal observers, and offer some convenient analytical methods for analyzing performance assessment.

2.5. Perspective on the Ideal Observer and System Optimization

The ideal observer not only gives strategies to analyze data, but the ideal observer SNR also gives a quantitative prediction of system performance for detection tasks. This means that one may compare two system designs to select which is better for detection tasks - or even plot SNR_I as a function of some system parameter for purposes of optimizing system design.

Although the ideal observer has considerable power to influence system design, a few items of perspective should be added.

2.5.1. Human Observer Performance

Maximizing the ideal observer signal-to-noise ratio means packing as much task specific diagnostic information into the raw data as possible. Ideally this will also improve human observer performance, however, it should be noted that human observer performance is greatly influenced by the processing and display algorithms used. Thus maximizing the ideal observer performance does not always mean maximizing human observer performance. As noted earlier, however, differences between human and ideal observer performance suggest an opportunity for additional image processing research to close the performance gap.

2.5.2. Task Specificity

Ideal observer system optimization is very task specific. A system optimization curve for detection of low contrast lesions may be completely different from a system op-

timization curve for the task of discriminating between a fluid filled cyst and a high contrast hypo-echoic solid mass. Cautionary notes should thus be made about system optimization. One of the goals of our research is to influence clinicians and system designers to think about medical ultrasound in a task specific way. We envision different programmable imaging modes on a scanner - each consisting of hardware and software modifications that are optimal for each diagnostic sub-task.

2.5.3. Lumpy Backgrounds

Previous literature has demonstrated that object statistics can influence system optimization.[31] There is a widely cited example of this in single photon emission imaging.[32] The BKE ideal observer SNR, when optimized for detector size, predicted erroneously that the detector should be as large as possible to collect as many photons as possible, even at the cost of spatial resolution. The theoretical remedy was to include stochastic 'lumps' in the background. The structure of the background was such that more false positives were introduced. A resulting optimization curve then predicted that there was a tradeoff between spatial resolution and sensitivity that was in part dependent on the object statistics. We only partially address this issue for ultrasonic imaging. It is a topic that will require future work.

2.6. Detection Theory in Ultrasonic Imaging

2.6.1. Smith-Wagner Theory

Smith and Wagner,[28] [29]pioneered some work in ultrasonic detection theory that we build on in this dissertation. We give some background on their work so that later

we may discuss connections with our new results.

Their statistical models of speckle were borrowed in part from Goodman,[33] who modeled scattering of coherent light from an ensemble of scatterers as a circular Gaussian complex process. In a similar manner, narrowband ultrasound echo signals at a given echo time can be modeled as obeying circular Gaussian statistics. The amplitude of the complex echo signal, b, considered as the envelope detected B-mode signal can subsequently be modeled as a Rayleigh distribution

$$pdf(b) = \begin{cases} \frac{b}{\psi} exp\left[-\frac{b^2}{2\psi}\right] & \text{if } b \geq 0 \\ 0 & \text{if } b < 0 \end{cases} \tag{15}$$

and the parameter ψ depends on the object's mean-square scattering strength (backscattered intensity). A vector of independent B-scan measurements thus has a distribution characterized by a product of Raleigh distributions. The likelihoods functions for the signal present (+) and signal absent (-) hypothesis are given as

$$pdf(\mathbf{b}|\psi_+) = \prod_{i=1}^{M} \frac{b_i}{\psi_+} \exp\left(-\frac{b_i^2}{2\psi_+}\right) \tag{16}$$

and

$$pdf(\mathbf{b}|\psi_-) = \prod_{i=1}^{M} \frac{b_i}{\psi_-} \exp\left(-\frac{b_i^2}{2\psi_-}\right) \tag{17}$$

The assumption of statistically independent measurements is not quite accurate for ultrasonic RF echo signals due to speckle correlations. Smith and Wagner nevertheless proceed with the independence assumption, assuming that the data measurements represent independent speckle spots rather than pixels. In this dissertation we

consider pixel correlations in a broader analysis.

The log-likelihood ratio for the signals, assuming independent samples, is then

$$\lambda_{Log-likelihood} = log\left(\prod_{i=1}^{M}\frac{\psi_+}{\psi_-}exp\left(\frac{b_i^2}{2}\left(\frac{1}{\psi_+}-\frac{1}{\psi_-}\right)\right)\right) \tag{18}$$

With a monotonic transformation, and ignoring additive and multiplicative constants, a sufficient test statistic is therefore

$$\lambda = \sum_{i=1}^{M} b_i^2 \tag{19}$$

The ideal decision function involves squaring and summing the readings over the cells in the region of interest.

Noting that the test statistic is a sum of many independent stochastic processes, we may conclude that it is Gaussian distributed. The moments are given as:

$$\begin{aligned} \langle\lambda|\psi_+\rangle &= 2M\psi_+; \qquad \langle\lambda|\psi_-\rangle = 2M\psi_-; \\ var[\lambda|\psi_+] &= 4M\psi_+^2; \quad var[\lambda|\psi_-] = 4M\psi_-^2; \end{aligned} \tag{20}$$

thus, the difference between the means is

$$\Delta\mu = \langle\lambda|\psi_+\rangle - \langle\lambda|\psi_-\rangle = 2M[\psi_+ - \psi_-] \tag{21}$$

and the common standard deviation is

$$\sigma = 2M^{1/2}[\psi_+^2 + \psi_-^2]^{1/2}. \tag{22}$$

Hence, the ideal observer signal-to-noise ratio for detection tasks, defined as the difference between the means over the common standard deviation is given as:

$$SNR_I = M^{1/2}\frac{\psi_+ - \psi_-}{\psi_1^2 + \psi_2^2}^{1/2} \tag{23}$$

This can be decomposed as follows:

$$SNR_I^2 = \frac{A}{S_c} SNR_o^2 C^2, \tag{24}$$

where A is the area of the lesion, S_c is the speckle correlation size,[29,28] SNR_o is pixel signal-to-noise ratio, equal to 1.91 for a Rayleigh process, and C is the image contrast, defined as $(\psi_+ - \psi_-)/\overline{\psi}$. This equation tells us that more independent speckle spots per target area yields better lesion detectability, consequently high spatial resolution improves our ability to detect even large low-contrast lesions. Detection is also easier for larger object contrast.

Like the Rose and Wagner-Brown models, this expression is important in that it connects the engineering features normally associated with image quality rigorously with task performance. The framework, however, was based on a number of stringent assumptions - some of which we hope to relax in this dissertation: (1) linear shift-invariance (2) large area lesions (3) low contrast lesions (4) independent samples (5) no electronic noise. Additionally, the authors analysis was based on Rayleigh densities of independent envelope detected samples rather than correlated RF signals. On the last point, note that envelope detection is a form of processing and that processing can either preserve or lose but never add diagnostic information. B-mode processing is in many cases not optimal, and that significant amounts of information in the phase of the signal can be useful for image reconstruction - especially in the case of phase-encoded waveforms.

2.6.2. Insana-Hall

Besides contributing to the interpretation of the Smith-Wagner theory above, Insana and Hall[34] performed human observer experiments to measure how close humans are at detection of low contrast lesions in B-mode images compared with the ideal observer. To do so they simulated pairs of ultrasonic image realizations (one image with a lesion and the other one without) using Monte Carlo simulations, and had human observers choose which of the images in the pair contained the lesion. Over a large ensemble of images, they found that human performance efficiency for the parameters they considered was about 50-60 percent.

Work done by Abbey and ourselves,[35] as discussed in future chapters, has shown that by including phase information, ideal observer performance is significantly better than Smith and Wagner predicted. Thus human observers are significantly less efficient at detecting important diagnostic image features of breast cancers which suggest a role for additional image processing techniques that make the information more accessible.

2.6.3. Clarkson-Barrett

Clarkson and Barrett[36] discuss the ideal observer for a number of different detection tasks and probability models. One of the models they discuss is the ideal observer for multivariate normal signals where the signal is in the variance. This is precisely the problem we wish to consider for ultrasound detection tasks. Their analysis was nevertheless complex and amenable to neither computation nor intuition, which we

very much desire. Their work is nevertheless important, as it is a rigorous treatment involving correlated data. We summarize the salient features of their contribution in a later chapter in the dissertation as we will make significant connections with their results.

2.7. Need for a New Model

Current models of detection performance in ultrasonic imaging have advanced the field of research but not pushed the field to its fundamental limits. This is the goal of our research. Our goal is to build on the contributions of Smith-Wagner, Clarkson-Barrett, and others to develop an ideal observer model where the ideal observer has access to the raw RF echo signals rather than the envelope-detected (lossy-processed) image data. We aim to glean not only image processing strategies, but also quantitative predictions and measurement techniques that can aid system designers in building the best possible system for important clinical tasks.

We begin by re-evaluating the way we model ultrasound systems and echo signal statistics. We then develop ideal observer models, and then illustrate a new imaging technique motivated by the theory.

3. LINEAR SYSTEM MODELS FOR ULTRASONIC IMAGING: APPLICATION TO SIGNAL STATISTICS

3.1. Introduction

The objective assessment of image quality in medical imaging systems is a topic of growing importance. Our ability to rigorously quantify the performance of clinically relevant tasks, such as identification of low-contrast lesions, can help us improve diagnostic performance and may influence standards of practice. Information theoretic approaches to performance assessment may help us understand fundamental limits of ultrasound systems, quantify information content of signals, and reveal optimum strategies for image processing. A key element of ultrasonic performance assessment is the computation of echo-signal covariance matrices that define likelihood functions. Current probabilistic models are valid only near the transducer focal region and in dense random media of point scatterers.[28,17] These models further assume the impulse response function is shift-invariant. Motivation for our work stems from the need to extend existing statistical models to more realistic shift-variant imaging systems and nonstationary random scattering media.

Pulse-echo imaging may be described as a linear system that obeys the imaging equation:

$$g(\mathbf{t}) = \int d\mathbf{x} h(\mathbf{x}, \mathbf{t}) \gamma(\mathbf{x}) + n(\mathbf{t}) = \mathcal{H}\gamma(\mathbf{x}) + n(\mathbf{t}), \tag{25}$$

where g is the radio-frequency (RF) echo signal consisting of a set of A-scan lines. The function h is the pulse-echo spatio-temporal impulse response, which is generally

shift variant. $\mathcal{H}$ is the linear operator associated with $h(\mathbf{x}, t)$. γ is called the object function (often stochastic), and n is signal-independent zero-mean additive Gaussian noise. Vector $\mathbf{t}$, represents, for example, the axial and lateral (temporal) dimensions of the echo data, will span the data space and therefore can increase in dimension depending on the study (see Appendix 3.A for details). Points in object space, in contrast, are positioned according to the Cartesian spatial vector $\mathbf{x} = (x_1, x_2, x_3)$. The dimensionality of the domain of h is the dimensionality of the object space plus that of the data space. Thus, for a single A-scan line, h is a four-dimensional (4-D) function: one dimension of time and three of space. Eq. (25) describes how spatial variations in physical properties of the object that interacts with the sound waves are mapped into data. It will describe acoustic transmission and reception (including the effects of diffraction, attenuation, and scattering), electromechanical coupling of the transducer, and electronic processing. It is a linear system in the sense that $\mathcal{H}\{a\gamma_1 + b\gamma_2\} = a\mathcal{H}\{\gamma_1\} + b\mathcal{H}\{\gamma_2\}$ for any object functions γ_1 and γ_2, and for any constants a and b. Thus, (25) assumes linearity in the system electronics, in the scattering response, and in acoustic propagation. Partial motivation for writing the imaging equation in form (25) is the immediate connection to literature on objective assessment of image quality.[11,30,13]

Investigators frequently use linear systems approaches in ultrasound research, some of which will be reviewed here. The imaging equation (25) is a generalization of many such approaches in the literature. Tupholme[37] and Stepanishen[38–41] were among the first to model ultrasonic imaging systems using the impulse response ap-

proach; the system impulse response was viewed as a series of convolutions of various impulse response functions representing acoustic, electromechanical, and scattering influences. This description led to a representation of a linear system, in which the object function was studied one point scatterer at a time. Other researchers have considered γ as a continuum variable that depends on perturbations in density and compressibility.[42,43]

Jensen[44] developed a linear model for the propagation and scattering of ultrasound in tissue based on a time-dependent Green's function solution to the wave equation in heterogeneous media. This is similar to our time-independent Green's function approach. However, he expresses the imaging equation 25 in terms of a spatial convolution [,[44] equation 45], and ignoring attenuation and noise. Our results reduce to his if we also make the assumptions of shift-invariance, attenuation-free medium, and no noise. Moreover, we take the extra step to apply the linear system theory to signal statistics.

Walker and Trahey[45] present a k-space linear systems model of ultrasound systems that they use to calculate speckle variance and echo correlations. Their application of the Fraunhofer approximation limits the results to the focal and farfield regions of the transducer. Walker[46] extends his results to shift-variant systems, but avoids making the an important distinction between point-spread and spatial sensitivity functions. Our results provide a rigorous underpinning to the computation of spatially varying statistics. Walker and Trahey[45] suggest that wavefront phase curvature does not contribute to speckle correlations. This conclusion was conditioned upon

some assumptions that will be elucidated later on. We demonstrate situations in which wavefront curvature can indeed be quite important in the evaluation of speckle texture.

Our aim is to demonstrate how the imaging equation (25) provides a comprehensive representation of shift-variant systems applied to continuous scattering media. Our motivation is to form a bridge between ultrasound physics and the statistical image quality assessment literature. The approach uses a solution of the homogenous wave equation propagating in inhomogeneous media. We relate our model to existing theory and show that it is well suited for evaluation of the statistical properties of ultrasound signals and images

3.2. Radiofrequency Signals

Here we begin from fundamental physics and proceed to model the RF echo signal along a single A-scan line. Thus here the vector $\mathbf{t}$ may be considered one dimensional t. Assume that the imaging medium is heterogeneous in density ρ and compressibility κ on a scale smaller than the wavelength, although the average density ρ, average bulk compressibility κ, and speed of sound, $c = 1/\sqrt{\rho_o \kappa_o}$ are approximately constant macroscopically throughout the medium. With the small amplitude approximation, acoustic propagation is well described by the homogeneous wave equation for inho-

mogeneous media *,[47–49]

$$\nabla \cdot (\frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x},t)) - \kappa(\mathbf{x}) \frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} = 0 \tag{26}$$

where $p(x,t)$ is the acoustic pressure field. Here time is a scalar quantity representing pulse-echo time for a single A-scan line. Taken from the classic acoustics literature,[50] Insana and Brown[49] provide a time-independent solution to (26) using a weak scattering Green's function approach. In particular, they give an expression for the pressure scattered to observation points r from a distributed scattering media. We will base our linear systems model on their approach, and connect it to linear systems approaches presented by other investigators. Backscattered energy at frequency ω propagates to the transducer aperture where it is coherently summed to generate a net instantaneous force, $f_\omega(t)$. Using the plane impedance approximation $p = \rho c u$ relating pressure p to particle velocity u normal to the detector surface S, and integrating over S, we obtain[51]:

$$f_\omega(t) = \frac{1}{2} i \rho_o c k U(\omega) exp(i\omega t) \int dx \left\{ k^2 \left(\frac{\Delta\kappa(\mathbf{x})}{\kappa_o} \right) A^2(\mathbf{x},k) + \left(\frac{\Delta\rho(\mathbf{x})}{\rho(\mathbf{x})} \nabla A(\mathbf{x},t) \cdot \nabla A(\mathbf{x},t) \right) \right\} \tag{27}$$

Eq. (27) is described in Appendix B of this chapter. Here $\Delta\kappa$ and $\Delta\rho$ are spatial fluctuations in κ and ρ about their respective means, $U(\omega)$ is the transducer surface

*The wave equation is homogeneous because we assume the medium contains no sources or sinks. The medium is inhomogeneous on a micro-scale because we allow there to be scattering sites.

velocity amplitude on transmission,

$$A(\mathbf{x},k) \equiv \frac{1}{2\pi}\int_S dS\xi(\mathbf{r})\frac{exp(-ik|\mathbf{r}-\mathbf{x}|)}{|\mathbf{r}-\mathbf{x}|} = 2\int_S dS\xi(\mathbf{r})G(\mathbf{r}|\mathbf{x}) \tag{28}$$

is the velocity potential due to a unit amplitude sinusoidal excitation,[48] $\xi(\mathbf{r})$ is the transducer apodization function, which may be complex to allow for focusing, and $G(\mathbf{r}|\mathbf{x})$ in (28) is the Greens function that describes the propagation of scattered waves created at field point $\mathbf{x}$ to observation points $\mathbf{r}$ on the transducer surface, as labeled in Fig. 1. Note that an assumption of separability between apodization (transducer geometry) and the driving velocity is made. The magnitude of the wave-vector k can be complex to allow for attenuation; for local plane waves $k = \omega/c - i\alpha(\omega)$, where α is the frequency dependent attenuation coefficient. The middle expression in (28) is the well-known Rayleigh integral,[33] representing a linear superposition of excitations from elements of the transducer surface.

Eq. (27) shows that scattering occurs at locations in the object of density and compressibility variations. The pattern of scattered sound is different for these two types of perturbations because density variations interact with the spatial gradient of the acoustic field A while compressibility variations interact with the field itself. Two copies of the field weightings A and ∇A are required to represent transmission and reception.

A weighted superposition of forces at each frequency in the transmitted pulse gives the measured echo-signal voltage (neglecting noise):

$$g(t) = \int d\omega Y(\omega) f_\omega(t) \tag{29}$$

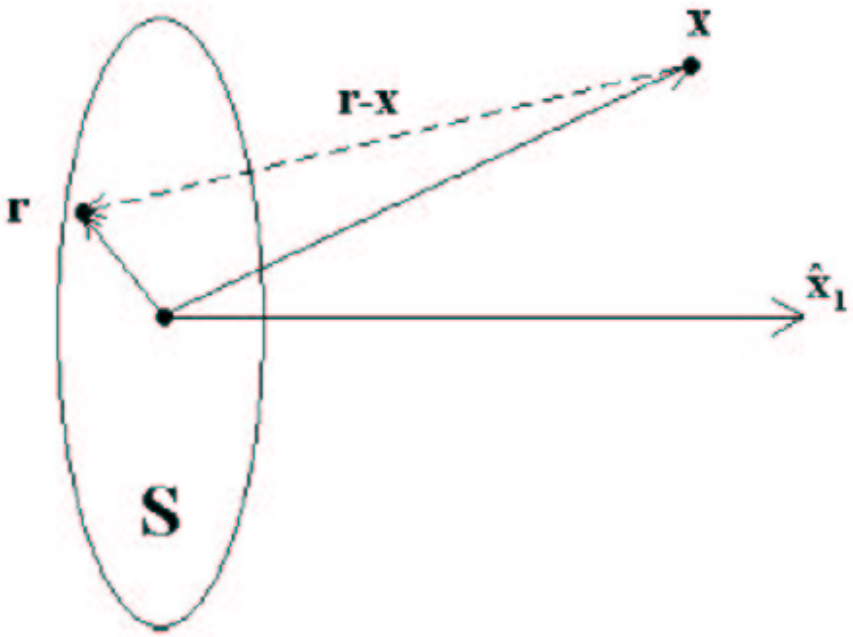

Figure 1. Illustration of the measurement geometry. Field and observation points are labeled **x** and **r**, respectively. Here a hat $\hat{\cdot}$ represents a unit vector.

$Y(\omega)$ is the complex electromechanical coupling coefficient of the transducer. Neglecting noise, the combination of (27) and (29) may be considered one form of linear system (25) operating on perturbations in density and compressibility. A more useful form of (29) may be obtained if one considers field points that are a distance greater than the aperture dimension[52] or media where perturbations in density contribute negligibly to the scattered field. In such cases, the pulse-echo spatiotemporal impulse response in (25) may be defined from (29) as:

$$h(\mathbf{x},t) = \int d\omega Y(\omega) \left[\frac{1}{2} i\rho_o ckU(\omega)\right] \kappa^2 A^2(\mathbf{x},k) exp(i\omega t) \tag{30}$$

and the object function γ in Eq. (25) represented as

$$\gamma(x) \equiv \left(\frac{\Delta\kappa(\mathbf{x})}{\kappa_o} - \frac{\Delta\rho(\mathbf{x})}{\rho(\mathbf{x})} \right) \cong -2\frac{\Delta z(\mathbf{x})}{z_o} \tag{31}$$

For local plane waves, $z(x) = \sqrt{\rho(x)\kappa(x)}$ is the acoustic impedance, z_o is the mean impedance, and $\Delta z(x) = z(x) - z_o$. The last expression in (31) holds only for small perturbations in density and compressibility.

Eq. (30) is an impulse response in the sense that it describes the response of the system to a point scatterer (a delta function in space). Pulsed excitation is incorporated by the superposition of frequency components initiated at transmission and filtered by the system. It is a straightforward exercise to show that (30) and (31) can be used to represent (25) in the time domain in terms of convolutions. The electromechanical impulse response is indicated by h_y. In the absence of dispersive attenuation, the transmit impulse response (which, by the principle of reciprocity is the same as the receive impulse response) is the velocity potential due to a temporal delta function excitation:

$$h_a(\mathbf{x}, t) = \frac{1}{2\pi} \int_S dS\xi(\mathbf{r}) \frac{\delta(t - |\mathbf{r} - \mathbf{x}|/c)}{|\mathbf{r} - \mathbf{x}|} = \mathcal{F}_t^{-1}\{A(\mathbf{x}, k)\}. \tag{32}$$

The last form reminds us that h_a and A are temporal Fourier transform pairs. Consequently, the echo signal may be expressed as:

$$g(t) = \int_V d\mathbf{x} \left[\left\{ h_y(t) * u(t) * \left(\frac{\partial}{\partial t} h_a(\mathbf{x}, t) \right) \right\} * s(\mathbf{x}, t) \right] + n(t) \tag{33}$$

Here the convolutions are over the time variable, u is the surface velocity, and s is the scattering response such that $s(t, \mathbf{x}) = \mathcal{F}_t^{-1}\{\gamma(x)k^2(\omega)\}$, where $\mathcal{F}_t^{-1}$ is the

inverse temporal Fourier transform operator. From the properties of convolutions, the temporal derivative can be moved to act upon any of the other functions. If the transmit and receive apertures are not the same, (33) should involve separate transmit and receive impulse response functions. With s taken as a point scatterer, and neglecting the k^2 filtering action of scattering, a form of (33) was derived with a different approach by Tupholme[37] and Stepanishen[41,38,39] and is the basis of a number of simulation packages and papers on linear systems.[53]

The spatiotemporal impulse response (30) may be written in terms of temporal convolutions:

$$h(\mathbf{x},t) = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left\{h_y(t) * h_y(t) * v(t) * \frac{\partial}{\partial t}h_a(\mathbf{x},t) * h_a(\mathbf{x},t)\right\} \tag{34}$$

where the surface velocity has been written as a convolution $u(t) = h_y(t) * v(t)$ between the transducer electromechanical impulse response and the driving voltage. Note that the driving voltage need not be an impulse to call h an impulse response function. It is an impulse response in the sense that a spatial impulse $\gamma(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_o)$ gives a noise-averaged linear response $h(\mathbf{x}_o, t)$. Also note that, although references (27) through (30) allow for dispersive attenuation by allowing a complex wave-vector, (**??**) do not, although one could easily add a plane-wave type of attenuation by convolving (34) in time with a position-dependent, pulse-echo attenuation filter. A practical alternative to the time-domain impulse response technique for evaluating (27) (30) computationally is the spatial-frequency domain angular spectrum method,[33] [54] that includes direction-dependent attenuation.

Like others[42–44] our approach describes a spatially distributed scattering function as a continuous distribution of point scatterers, and it attaches physical meaning to s by connecting it to variations in density and compressibility. Contrary to (33), in (30) and (34) we choose to couple the k^2 filtering action of scattering (object properties) with acoustic and electromechanical impulse response functions (system properties). It should be clear that there is some flexibility in defining which factors are associated with the object and which with the system. That decision depends on the application, and this flexibility is considered a strength of the linear systems approach.

3.3. Point-Spread and Spatial Sensitivity Functions

As discussed in Appendix A, we may consider the data space variable $\mathbf{t}$ to be a vector representing, for example, the axial and lateral dimensions of a B-mode image. With this interpretation, various samples in the data may be the result of different spatiotemporal, impulse-response functions, i.e., shift variance. Imaging systems are often studied using point-spread functions. An ultrasonic point-spread function (psf) is obtained from the ensemble of echoes recorded from one stationary point scatterer. The psf is a data-space slice though the higher dimensional impulse response $h(\mathbf{x}, \mathbf{t})$. Point-spread functions characterize the spatial resolution of the imaging system. Consider the psf representing noisy RF data from a point scatterer fixed at location x_o:

$$psf_{\mathbf{x}_o}(t) = \langle \int d\mathbf{x} h(\mathbf{x}, \mathbf{t}) \delta(\mathbf{x} - \mathbf{x}_o) \rangle_n = \int d\mathbf{x} h(\mathbf{x}, \mathbf{t}) \delta(\mathbf{x} - \mathbf{x}_o) = h(\mathbf{t}|\mathbf{x}_o) \qquad (35)$$

Here, the notation $\langle \cdot \rangle_n$ represents an ensemble average over noise variations in the stochastic quantity n. When the noise level is sufficiently low, averaging can be ne-

glected. The spatiotemporal impulse response h at time $\mathbf{t}$ and position $\mathbf{x}_o$ is equivalent to the point-spread function at sample time t obtained when scanning a point source at position $\mathbf{x}_o$.

Shown in Figs. 2(a)(c) are RF data matrices illustrating the psfs $g(t_1, t_2) = h(t_2, t_1|\mathbf{x}_o)$ generated using Field II [24] by scanning a point scatterer placed at three different positions in the field of a linear array transducer: nearfield, focal zone, and farfield. Fig. (2) was formed by pulsing and receiving along a number of parallel lateral A-scan lines. With this scanning configuration, at any given depth, the point-spread function is conveniently shift-invariant in the lateral (vertical) direction. At first glance, the curvature of the psf wavefronts may seem counterintuitive. One might expect, for example, a transmitted wavefront to be concave in the nearfield because concave focusing is being used. Convex nearfield curvature is observed because the pulse-echo transit time is smaller when a (concave) transducers edge is nearer in pulse-echo transit time to the point scatterer than when the scatterer is centered with respect to the transducer. Also note that the phase oscillations in the psf maintain a constant frequency along each A-scan line. This is why the width of the psf in the direction normal to the wavefront seems to narrow at the edges. The low magnitude X-shape components in the psf are focusing flaws and include edge waves.[55]

Of significance for modeling signal statistics is the spatial distribution of h for a fixed time i.e., $h(\mathbf{x}|\mathbf{t})$, which we refer to as the spatial sensitivity function (ssf). This is to be contrasted with the psf that describes temporal rather than spatial variation. The ssf is an object space slice though $h(\mathbf{x}, \mathbf{t})$ that describes how the impulse response

for a fixed-measurement time changes as the point scatterer is moved in space. An alternative perspective is that, for each data sample, the ssf describes how sensitive the imaging system is to point scatterer positions located throughout the object. Figs. 2(d)-(f) display spatial sensitivity functions corresponding to nearfield, focal region, and farfield pulse-echo times. They were obtained by generating a sequence of psfs: $h(\mathbf{t}|\mathbf{x}), h(\mathbf{t}|\mathbf{x}+\Delta\mathbf{x}), ..., h(\mathbf{t}|\mathbf{x}+N\Delta\mathbf{x})$, where $\Delta\mathbf{x}$ represents an increment in the axial direction. Each psf was sampled along a particular pulse-echo axial time point $\mathbf{t}$, and the resulting lateral scan vectors corresponding to each incremental position were assembled into a spatial sensitivity image. It should be emphasized that, for a given time in data space, a ssf is a 3-D function of space, and we are only displaying a 2-D slice through this distribution. The dimensionality of the psf depends on the dimensionality of the data space: for 2-D imaging the psf will be 2-D and for 3-D imaging the psf will be 3-D.

Spatial sensitivity functions are essential for studying first- and second-order moments of samples or pairs of samples in the data set. For example, to study the covariance between samples recorded at $\mathbf{t}$ and $\mathbf{t}'$, and modeling the object function as a zero-mean, unit-variance, white Gaussian random variable, we need to compute:

$$\begin{aligned}\langle g(\mathbf{t})g(\mathbf{t}')\rangle_{n,\gamma} &= \int\int d\mathbf{x}d\mathbf{x}'\langle\gamma(\mathbf{x})\gamma(\mathbf{x}')\rangle_\gamma h(\mathbf{x}|\mathbf{t})h(\mathbf{x}'|\mathbf{t}') + \langle n(\mathbf{t})n(\mathbf{t}')\rangle_n \\ &= \int d\mathbf{x}h(\mathbf{x}|\mathbf{t})h(\mathbf{x}|\mathbf{t}') + \sigma_n^2\delta(\mathbf{t}-\mathbf{t}') \qquad (36)\end{aligned}$$

which requires spatial sensitivity functions. Techniques for calculating statistical moments for shift-invariant and shift-variant systems were presented previously,[46]

but these contributions presumed the point-spread function as the critical quantity. Later on, we emphasize that a local shift-invariance approximation is needed to use point-spread functions in place of sensitivity functions.

3.4. In-Phase and Quadrature Signals

Real RF echo signals are expressed as complex functions when the task is to estimate the amplitude or phase of the signals separately; for example, to estimate blood velocity.[56] The in-phase (I) and quadrature (Q) decomposition is expressed by

$$g(t) = g_I(t)cos(\omega_o t) - g_Q(t)sin(\omega_o t) = Re\left\{exp(i\omega_o t)[g_I(t) + ig_Q(t)]\right\} \qquad (37)$$

The quantity in curly brackets is the analytic signal $g(t) + i\tilde{g}(t)$, where the tilde represents the Hilbert transform of $g(t)$. Because in-phase and quadrature components g_I and g_Q are baseband signals, they can typically be sampled at a much lower rate than the RF echo signal. This representation saves data transfer and computational times while preserving information content in the signal. Because of its common use, it is desirable to model the $[I, Q]$ decomposition with a linear systems approach. To this end, the echo signal may be written as:

$$g(t) = \int d\mathbf{x}\gamma(\mathbf{x})e^{i\omega_o t}\mathcal{F}^{-1}\left\{H(\omega - \omega_o|\mathbf{x})\right\} + n(t) = e^{i\omega_o t}\int d\mathbf{x}\gamma(\mathbf{x})h_o(\mathbf{x}, t) + n(t) \quad (38)$$

where H is the frequency response of h and h_o is the demodulated spatiotemporal impulse response and in general is complex. The in-phase and quadrature components of an echo signal thus may be modeled from (25) as:

$$g_I(t) \;\; = \;\; \int d\mathbf{x}\gamma(\mathbf{x})h_I(\mathbf{x}, t) + n_I(t)$$

$$g_Q(t) \quad = \quad \int d\mathbf{x}\gamma(\mathbf{x})h_Q(\mathbf{x},t) + n_Q(t) \tag{39}$$

Here n_I and n_Q are assumed to be independent Gaussian noise processes. The functions h_I and h_Q are the real and imaginary parts of h_o, and will be referred to as in-phase and quadrature spatiotemporal impulse response functions, respectively. Shown in Fig. 3 are the images of h_I and h_Q, visualized in the spatial and temporal domains; in other words, the in-phase and quadrature point-spread and spatial sensitivity functions. Note that the I and Q psfs Figs. 3(a) and (b) contain no axial phase oscillations, as these temporal variations have been demodulated. The lateral banded pattern occurs because of the pulse-echo wavefront curvature. A changing temporal axial phase thus shifts energy between the in-phase and quadrature channels. The I and Q ssfs of Figs. 3(c) and (d) are markedly different than the I and Q psfs in that the former exhibit axial spatial phase ($exp(ikx)$) oscillations. The in-phase and quadrature ssfs are 90^o out of phase and, therefore, those echo signals are uncorrelated, which we will find later is a very important feature for computing the ideal observer test statistic.

By replacing $h(\mathbf{x}|\mathbf{t})$ with $h_I(\mathbf{x}|\mathbf{t})$ in 36, the covariance of the in-phase signal at times $\mathbf{t}$ and $\mathbf{t}'$ may be computed. Current theory, valid only in the focal region, suggests that the statistical properties of fully developed speckle depend only on the overall shape of the pulse envelope,[28] [17] and is independent of phase information. To extend this statistical analysis outside of the focal zone, 36 suggests that the precise phase of h or $[h_I, h_Q]$ is required to accurately model the statistical moments

of the echo data $\mathbf{g}$. Statistical descriptions of echoes throughout the field is the main application of this linear systems analysis.

3.5. B-Mode Signals

B-mode imaging systems display the envelope of the RF echo signal obtained from the magnitude of the $[I, Q]$ data or equivalently the magnitude of the analytic signal:

$$g_e(t) = \sqrt{g_I^2(t) + g_Q^2(t)} = \sqrt{g^2(t) + \tilde{g}^2(t)} \tag{40}$$

Envelope detection is inherently nonlinear, thus linear systems analysis of ultrasound systems must be applied to either the RF or $[I, Q]$ data. The envelope of h, which we denote $h_e(x, t)$, is important for modeling focal-region statistics of the B-mode image signals,[28] ,[17] and may be written as $h_e = |h_o| = \sqrt{h_I^2 + h_Q^2} = \sqrt{h^2 + \tilde{h}^2}$. Because B-mode processing discards phase information, it is natural to wonder whether envelope detection is the optimal strategy for data visualization and, if not, what is the optimal strategy? Current research aims to address this question in the context of the image quality assessment literature.[13]

3.6. The Local Shift-Invariance Approximation

As shown in Appendix 3.A, 2-D image reconstruction maps the temporal coordinates of the echo data $g(t_1, t_2)$ into spatial coordinates associated with image pixels, $\widehat{\gamma}(\mathbf{x}_1, \mathbf{x}_2)$. Furthermore, the spatiotemporal coordinate transformation is linear, $\mathbf{t} = \mathbf{B}\mathbf{x}$, and therefore interesting relationships exist between point-spread and spatial

sensitivity functions when linear shift-invariance (LSIV) holds. If we find that

$$h(\mathbf{x},\mathbf{t}) \cong h(\mathbf{x}+\Delta\mathbf{x},\mathbf{t}+\mathbf{B}\Delta\mathbf{x}) \tag{41}$$

over some range of $\Delta\mathbf{x}$, then we say that h is LSIV for values of $\mathbf{x}$ where (41) holds. This region is often labeled isoplanatic. For such regions, h may be written as a function of one vector variable:

$$h(\mathbf{x},\mathbf{t}) = h(\mathbf{t}-\mathbf{B}\mathbf{x}) \tag{42}$$

thus,

$$ssf(\mathbf{x}) = h(\mathbf{t}_o - \mathbf{B}\mathbf{x}|\mathbf{t}_o)|_{\mathbf{t}_o=\mathbf{B}\mathbf{x}_o} = h(-[\mathbf{B}(\mathbf{x}-\mathbf{x}_o)]|\mathbf{x}_o) = psf(-(\mathbf{t}-\mathbf{t}_o)) \tag{43}$$

This means that the ssf, a function of $\mathbf{x}$, is a copy of the psf time-reversed about the time point $\mathbf{t}_o = \mathbf{B}\mathbf{x}_o$. This is why there appears to be reflective symmetry between the psfs and ssfs shown in Fig. 2. It is emphasized that the symmetry is only approximate.

Eq. (43) is not valid and the system is not locally LSIV when amplitude profiles or wavefront curvatures vary significantly with position, as is the situation for strongly focused transducers. Fig. 4 shows farfield ssf and time-reversed psf A-scan lines Fig. 4(a) on axis and Fig. 4(b) 7.6 mm off axis. The phase mismatch off axis may be explained by the changing curvature of the wavefront over an assumed isoplanetic patch. The approximation is much better on axis [Fig. 4(a)] or in the focal region, where wavefront curvature is minimal.

In an isoplanetic region, the shape of the point-spread or spatial sensitivity functions do not change significantly. One way of assessing the approximate size of an isoplanetic patch is to plot the normalized correlation coefficient between different sensitivity or point-spread functions as a function of temporal or spatial separation. Because it is often easier to compute the psf rather than the ssf, we choose to assess how the psf changes as a function of distance. More precisely, let $p_1(\mathbf{t}) = h(\mathbf{t}_1|\mathbf{x})$ and $p_2(\mathbf{t}) = h(\mathbf{t}|\mathbf{x}_2 = \mathbf{x}_1 + \Delta\mathbf{x})$ be point-spread functions for spatial (object space) points $\mathbf{x}_1$ and $\mathbf{x}_2$. Then the normalized correlation coefficient for assessing shift invariance is:

$$C_{p_1p_2}(\Delta\mathbf{x}) = \frac{\int p_1(\mathbf{t})p_2(\mathbf{t} - \mathbf{B}\Delta\mathbf{x})d\mathbf{t}}{\int p_1(\mathbf{t})d\mathbf{t} \int p_2(\mathbf{t})d\mathbf{t}} \tag{44}$$

$C_{p_1p_2}(\Delta\mathbf{x})$ is a measure of the similarity between overlaid point-spread functions. By plotting $C_{p_1p_2}(\Delta\mathbf{x})$ as a function of lag Δx, the degree of shift invariance can be assessed. When $C_{p_1p_2}(\Delta\mathbf{x})$ falls below a threshold value (e.g., 90%) we say that shift invariance fails. The supra-threshold region defines the extent of the isoplanetic patch. To relate the point-spread and spatial sensitivity functions as in (43), the temporal shape of the point-spread functions should not change as the scatterer location is moved spatially over the extent of the scattering volume associated with a given instant of time. Alternatively, the size of the isoplanetic patch should be larger than the temporal extent of the point-spread function itself, as seen in Fig. 5. The LSIV approximation holds across lateral scan lines for linear arrays, except when the beam is electronically steered or near the margins in which the aperture is reduced. The isoplanetic regions will be small along the beam axis for fixed-focus sources in which

the wavefront curvature varies. Outside the focal region, highly curved, extensive ssf wavefronts may exhibit enough axial spatial extent to exceed the size of an isoplanetic patch. In this case, (43) does not hold and the symmetry between ssf and psf is lost. The 90% threshold isoplanetic patch for the farfield psf of Figs. 2(c) and (f), extended from 88.8 to 91.3 mm. Significant energy exists beyond this region; hence the lack of symmetry between the psf and ssf as seen in Fig. 4. Systems with dynamic focusing and aperture growth which includes all commercial systems, as well as systems that use synthetic aperture approaches may be expected to have fairly large isoplanetic regions.

The local LSIV of isoplanetic regions fails for in-phase and quadrature signals. As is evident in Fig. 3, the I and Q ssfs are not approximated by time reversed I and Q psfs. However, the I and Q ssfs can be represented as a phase-shifted version of the RF ssfor the time-reversed RF psf. Multiplying the phase factor $exp(i\omega_o t)$ and the analytic ssf signal (approximated by the time-reversed psf whose argument has been converted to a spatial coordinate) $h(\mathbf{x}|t)+i\tilde{h}(\mathbf{x}|t)$ and taking the real part of the result, the in-phase ssf may be obtained. The quadrature ssf may be obtained by taking the imaginary part of the product. Thus, with a local shift-invariance approximation, the RF point-spread functions may be used to compute statistical moments of the in-phase and quadrature signals without computing sensitivity functions. This may be useful when it is more natural to calculate psfs than ssfs, as is the case with many simulation packages.

When an isoplanetic region is larger than the sensitivity function, the LSIV ap-

proximation holds and (25) reduces to a spatial convolution, e.g., $\overline{g(\mathbf{x})} = [h * \gamma](\mathbf{x})$. The LSIV assumption allows us to express g and $\overline{g} = \langle g \rangle_n$ as functions of either $\mathbf{x}$ or $\mathbf{t}$. Specifically, $g(t(\mathbf{x})) = g(\mathbf{B}\mathbf{x})$. Under the isoplanetic assumption, our results reduce to those of Jensen [Ref.,[44] equation 45]. Walker and Trahey[45] chose to represent the LSIV imaging equation in frequency or k-space. In the Fraunhofer regime, this has a particularly elegant interpretation as the field profile is simply the Fourier transform of the aperture.

Although the validity of the local LSIV assumption for the RF signal is largely restricted to the focal region of array transducers, the assumption nevertheless is useful for analysis and interpretation. Let us define $\breve{h}$ as the pulse-echo impulse response, neglecting the quadratic frequency dependence of scattering [i.e., $\breve{h}$ is the LSIV approximation to the quantity in curly brackets in (33)]. This approximation allows the noise-averaged linear system to be written in the spatial frequency domain as[2]:

$$\overline{G}(\mathbf{k}) = \mathcal{F}_{\mathbf{x}}\left\{\overline{g(\mathbf{x})}\right\} = H(\mathbf{k})\Gamma(\mathbf{k}) = k_1^2 \breve{H}(\mathbf{k})\Gamma(\mathbf{k}) = \breve{H}(\mathbf{k})\Phi(\mathbf{k}) \tag{45}$$

where $\mathbf{k} = (k_1, k_2, k_3)$ is the spatial angular-frequency vector (conjugate to $\mathbf{x}$), and H, $\breve{H}$, and Γ are the spatial Fourier transforms of h, $\breve{h}$, and γ, respectively. The scattering amplitude is defined as:

$$\Phi(k) = -\frac{1}{4\pi} k_1^2 \Gamma(k) = \frac{1}{2\pi z_o} \mathcal{F}_{\mathbf{x}} \left\{ \frac{\partial^2 z(\mathbf{x})}{\partial x_1^2} \right\}, \tag{46}$$

where x_1 is the component of x normal to the transducer surface as in Fig. 1. A local plane wave approximation is required to write (45) and (46), and hence its validity is

restricted, in practice, to the focal region of ultrasound B-scans. The noise-averaged echo signal can now be written as,[2 57]:

$$g(x) = -2\left[\breve{h} * \frac{\partial^2}{\partial x_1^2}\left(\frac{z}{z_o}\right)\right](\mathbf{x}) = -2\left[h * \left(\frac{z}{z_o}\right)\right](\mathbf{x}) \tag{47}$$

The first equality in (47) illustrates that sound is scattered whenever the second derivative of the relative acoustic impedance (in the direction of transmission) is nonzero. The second equality allows one to consider the acoustic impedance itself as the object function. The spatiotemporal impulse response function, h, thus acts as a spatial frequency filter on the object function $z(x)$. The spatial frequency response of this filter is described by H, which is the k-space picture of h.[45] Insana and Cook[58] give a useful LSIV approximation for a beam using a Gaussian signal model. They show that h acts as a bandpass filter of $z(x)$ in the axial direction, and as a lowpass filter of $z(x)$ in the lateral direction. The bandpass nature of h in the scanning direction is due to modulation about a carrier frequency, and the low-pass character in the lateral direction is due to the (unmodulated) taper of the beam profile.

3.7. Signal Statistics

The linear systems framework developed here and by others provides a starting point for analysis of signal statistics. Eq. (36) illustrates one way to compute the covariance of the RF signal. With a local isoplanatic assumption, the point-spread function can be used instead of the spatial sensitivity, and our results match those of Walker and Trahey,[45] with one important exception. They predicted that wavefront curvature plays no role in the RF signal covariance between two signals when the respective

apertures are coplanar. Their conclusion was based on stringent assumptions, including the Fraunhofer approximation and an assumption that phase curvature has no lateral positional dependence over some region of interest that does not extend too far off axis. Explicitly, for spatial positions $\mathbf{x}$ far from the aperture $x_1 \gg k|\mathbf{r}|^2_{max}/2$ and not too far off axis, so that $\cos(\mathbf{x}_1, \mathbf{r}-\mathbf{x}) = 1$, the continuous wave (CW) Fresnel approximation of (28) due to a sinusoidal excitation:

$$A(\mathbf{x},k) = \frac{exp(ikx_1)}{i\lambda x_1} exp\left[i\frac{k}{2x_1}(x_2^2+x_3^2)\right]\int_{-\infty}^{\infty}\left\{\xi(\mathbf{r})exp\left[-i\frac{k}{2x_1}|\mathbf{r}|^2\right]\right\}exp\left[-i\frac{2\pi}{\lambda x_1}(\mathbf{x}\cdot\mathbf{r})\right]dS \tag{48}$$

reduces to a Fraunhofer approximation:

$$A(\mathbf{x},k) \simeq \frac{exp(ikx_1)}{i\lambda x_1} exp\left[i\frac{k}{2x_1}(x_2^2+x_3^2)\right]\int_{-\infty}^{\infty}\xi(\mathbf{r})exp\left[-i\frac{2\pi}{\lambda x_1}(\mathbf{x}\cdot\mathbf{r})\right]dS \tag{49}$$

This is because the quadratic phase term $exp\left[-i\frac{k}{2x_1}|\mathbf{r}|^2\right]$ in the Fresnel expression is approximately unity over unfocused aperture faces for points x_1 far enough away. When a focused aperture is used, the apodization ξ may be considered a complex quantity. Its purpose is to negate the $exp\left[-i\frac{k}{2x_1}|\mathbf{r}|^2\right]$ phase term in the Fresnel integral and, hence, simulate the Fraunhofer region at a much closer range. The quadratic phase argument will be negligible at some depth of field about the focus of the transducer. In pre- or post-focal regions, where x_1 is comparable to or less than $k|\mathbf{r}|^2_{max}/2$, however, the phase may be significant and the Fraunhofer approximation may fail. In these situations, phase curvature is dependent on lateral position, and the approximations of Walker and Trahey[46] do not hold. Their conclusion is, for example, applicable to the focal region of focused transducers in which wavefront curvature is essentially

flat and farfield regions of unfocused transducers (as long as energy is concentrated in a region not too far off axis). We give an example of how wavefront curvature can play a very important role in predicting speckle texture for focused transducers in pre- and post-focal regions. Consider an electronically swept linear array system with fixed focus and shift invariance in the lateral direction. We are interested in the correlation lengths of speckle along lateral scan lines. In this case the covariance (36) depends only on the difference $(\mathbf{t} - \mathbf{t}')$, and thus is simply a cross correlation. Eq. (36) tells us that, to compute the speckle cross correlation, one should slide the sensitivity map laterally, then multiply with a copy of itself, and integrate.

In the near- and farfield, wavefront curvature is significant, and this procedure of shifting and multiplying will result in complex phase interference patterns that will quickly integrate to zero. Fig. 6 shows the predicted correlation coefficients of in-phase speckle texture in the lateral direction for curved and flat wavefront sensitivity functions along with the corresponding erroneous predictions that use in-phase psfs instead of ssfs, phase, and neglect phase or assume a flat wavefront field of equal shape and scattering volume. Speckle patterns corresponding to the curved wavefront field decorrelate more quickly than does the flat wavefront field. This is one reason why speckle texture appears fine in the near and farfield regions of ultrasound B-scans even though the pulse energy is spread out very broadly. A related connection to the literature concerns phase aberration. Speckle has been observed to be broadest whenever aberrations are least,[59] .[60] Phase aberrations cause irregularities in phase fronts that induce rapid decorrelation even though the pulse volumes are very large.

The isoplanetic approximation to the lateral correlation coefficient of Fig. 6 (not shown) for lateral speckle correlation is very good (less than 1% maximum deviation) because lateral shift invariance is assumed. Fig. 7 shows correlations in the axial direction, and the corresponding isoplanetic approximation. In the axial case the isoplanetic assumption deviates significantly (almost 20% at 1-mm lag) from the full shift-variant computation. Better agreement may be expected for shorter pulses because the axial correlation distance will be shorter relative to the size of the isoplanetic patch. For simplicity, in all the examples shown, we imagine that we are imaging a 2-D planar distribution of scatterers, so that we can assume the ssf is a 2-D instead of a 3-D function.

If the system is locally shift invariant, and the object can be modeled by a zero mean, wide sense stationary (WSS) random process, then the covariance matrices are completely characterized by autocorrelation functions. In the continuous sampling limit, the eigenvalues of the covariance matrix are simply the power spectrum. Thus,

$$\begin{aligned} K(x_1, x_1 + \Delta\mathbf{x}) &= R(\Delta\mathbf{x}) \\ &= h(\Delta\mathbf{x}) * R_\gamma(\Delta\mathbf{x}) * h(-\Delta\mathbf{x}) + R_n(\Delta\mathbf{x}) \\ &\leftrightarrow |H(\mathbf{k})|^2 S_\gamma(\mathbf{k}) + NPS(\mathbf{k}) \end{aligned} \tag{50}$$

where $R_\gamma(\Delta\mathbf{x})$ and $S_\gamma(\mathbf{k})$ are the autocorrelation and power-spectral representations of the object function, respectively, and $R_n(\Delta\mathbf{x})$ and $NPS(\mathbf{k})$ are the autocorrelation and power spectrum of the noise, respectively.

One curiosity is that the covariance itself depends on wavefront curvature, yet from

(26), only the magnitude of $H(\mathbf{k})$ has a role in speckle properties. The conclusion is that curved wavefront ssfs have spatial phase curvature information in $|H(\mathbf{k})|$.

Intriguingly, the lateral spatial frequency bandwidth [Fig. 4(d)] of the large curved wavefront sensitivity in Fig. 4(c) is broader than the bandwidth [Fig. 4(b)] of the smaller focal region sensitivity of Fig. 4(a). Specifically, the effective lateral bandwidth of Fig. 4(d) as defined by Bendat and Piersol,[61] and evaluated at the axial frequency maximum was 74% broader than that of Fig. 4(b). Thus, from an information-theoretic point of view, the curved wavefront ssf has potentially more spatial resolution available, even though the pulse volume is considerably larger! This can be better understood by realizing that a slice through the sensitivity function in the lateral direction looks like a chirp function. A matched or mismatched compression filter could potentially be applied to regain spatial resolution while improving signal to noise, similar to current coded excitation schemes.[62] A fascinating possibility for designing ultrasound systems could be the introduction of an engineered, coded excitation in the lateral direction of the beam, similar to work that has been done for 2-D velocity estimation using array transducers,[63] .[64]

As a final illustration of the application of this work, we give an example of how shift-variant systems may exhibit statistical coupling of in-phase and quadrature signals along axial scan lines. Most current literature assumes independence of I and Q channels. It is well-known that the on-axis acoustic response consists of direct and edge waves, and that these contributions will have a position dependent phase delay. Keeping this in mind, now consider computing the covariance between in-phase and

quadrature signals at various points of time. Eq. (36) tells us that we should multiply sensitivity functions for the I and Q signals then integrate. This integration will not tend to zero if portions of the in-phase signal are not exactly 90^o. out of phase with the quadrature, as demonstrated in Fig. 9. Statistical independence between I and Q signals can be expected in the focal zone of unapodized transducers as direct and edge waves are approximately superimposed with a constant phase relationship. Understanding these edge-wave effects may be important for understanding or eliminating unwanted image correlations, both from a system design point of view and an image processing perspective. The degree of statistical dependence between I and Q channels will be diminished greatly with decreasing edge-wave amplitudes. Consequently, apodization may significantly reduce edge wave amplitudes and abate $I - Q$ correlation.

3.8. Conclusions

A linear systems framework based on a solution to the wave equation for inhomogeneous media has been presented that is similar to others.[42–44] With certain approximations, our results reduce to a continuum extension of the Tupholme-Stepanishen theory.[37–41] Although their focus has been on point-spread functions, ours is on spatial sensitivity functions because of their value for predicting task performance. For each echo sample, the spatial sensitivity function reveals the sensitivity of the ultrasound system to each location in the object. These functions, along with their in-phase and quadrature counterparts are important in the computation of statistical

moments. When local shift-invariance holds (as is the case near the focal region of an array transducer), the spatial sensitivity functions are shown to be similar to point-spread functions, which are time reversed about the axial position of the scatterer location. This time-reversal property does not apply to the in-phase and quadrature distributions. The LSIV approximation together with a plane wave approximation has further been used to understand the system response as a spatial frequency band-pass filter of acoustic impedance $z(x)$ in the axial direction, and as a lowpass filter of $z(x)$ in the lateral direction. An equivalent interpretation suggests that ultrasound systems are sensitive only to the second derivative of $z(x)$ with respect to the scanning direction. In situations in which local shift invariance does not apply, the full shift-variant theory must be used. Shift-variant, in-phase and quadrature spatial sensitivity functions may be particularly important for the evaluation of statistical properties of demodulated or envelope-detected images of realistic ultrasound imaging systems. Such statistical characterization may be useful for quantifying image quality and for design of image processing algorithms. Our theory predicts that wavefront curvature and phase information contained within sensitivity functions is essential for explaining the complex speckle structure apparent in ultrasound images. Coupling between direct and edge waves introduces statistical correlations between I and Q channels. Also revealed is an opportunity to enhance near- and farfield spatial resolution by matched filtering unfocused beams. Ultimately this work will aid our understanding of fundamental performance limits, optimum image processing strategies, and quantitative image quality metrics for ultrasonography.

3.9. Appendix A

This appendix provides examples of object, data, and image function representations that may be useful in several ultrasonic applications of the linear imaging equation (25).

One common data structure for B-mode acquisition is shown in Fig. 10. A linear array generates beamformed echo sequences oriented in columns with elements $g(\mathbf{t}) = g(t_1, t_2) = g(\ell T, mLT)$. The L range samples are placed in columns corresponding to each A-line and assigned a discrete time stamp $t_1 = \ell T$, where $0 \leq \ell \leq L-1$ are integers and T is the temporal sampling interval. Columns are filled left to right with sequenced A-line recordings, where $t_2 = mLT, 0 = m = M-1$. The acquisition time is given by $t' = t_1 + t_2 = (\ell + mL)T$, and the integer ℓ is indexed completely between 0 and $L-1$ before m is incremented. In this example, sequentially acquired data are represented by a 2-D matrix of echo samples $g[\ell, m]$. It is often convenient to organize all the data from a scan-plane into a single LM-dimensional column vector $\mathbf{g}$ with elements $g[\ell + mL]$. Then from (25) we can write the imaging equation as a continuous-to-discrete transformation:

$$\mathbf{g} = \int dx h(\mathbf{x}, \mathbf{t}'[\ell, m])\gamma(\mathbf{x}) + \mathbf{n} = H\{\gamma(\mathbf{x})\} + \mathbf{n} \tag{51}$$

where $\mathcal{H}$ in the last form is a linear operator on γ that generates g.

To facilitate diagnostic interpretation, human observers require that echo locations be one-to-one with the corresponding object locations. Consequently, we apply a discrete-to-discrete reconstruction operator $\mathcal{O}$, viz., $\hat{\gamma} = \mathcal{O}g$, that converts echo data

into an image of the object, $\widehat{\gamma}$. Like $\mathbf{g}$, $\widehat{\gamma}$ is a vector of length LM. If $\mathcal{O}$ is linear, then[11]:

$$\widehat{\gamma} = \mathcal{O}\mathcal{H}\gamma(\mathbf{x}) + \mathcal{O}\mathbf{n} = \mathcal{S}\gamma(\mathbf{x}) + (\mathcal{O}\mathcal{H} - \mathcal{S})\gamma(\mathbf{x}) + O\mathbf{n} \tag{52}$$

where $\mathcal{S}$ is the sampling operator. The first term on the right side is the sampled object, the second term is the bias between the reconstructed image and sampled object, and the third term is the image noise. If the task is to obtain an image of the object nearest its true form, a superior linear imaging system minimizes the second and third terms. B-mode image reconstruction is nonlinear, so (52) does not apply. Nevertheless B-mode image reconstruction is straightforward. Essentially we take the envelope of the echo data, [see (40)] and convert temporal coordinates to spatial coordinates: $t_1 = \ell T \rightarrow \ell cT/2 = \ell \Delta X_1$, and $t_2 = mLT \rightarrow m\Delta X_2$, where c is the speed of sound and ΔX_1 and ΔX_2 are the vertical and horizontal pixel dimensions corresponding to the axial and lateral (pitch) spatial sampling intervals. As part of the reconstruction process, we normally convert the temporal axes of the data into spatial axes via the inverse of the coordinate transformation $\mathbf{t} = \mathbf{B}\mathbf{x}$, where, in the case of swept-scan 3-D B-mode imaging, B is a diagonal matrix. Specifically, the mapping is:

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} 2/c & & 0 \\ & T/\Delta X_2 & \\ 0 & & T/\Delta X_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{53}$$

To consider other ultrasonic techniques requires that we expand the dimension of the time vector $\mathbf{t}$. Doppler estimates use several A-line acquisitions at each lateral

transducer position. Fast-time (range) samples, identified above by the index ℓ, are repeatedly acquired Q times before indexing m to form a slow-time data set at times $t_o = qLT$.[56] The integer index is $0 \leq q \leq Q-1$, and we assume the temporal pulse-repetition interval is $T_p = LT$. Further, to add data for either 3-D, B-mode imaging or strain imaging, sequential scan planes are acquired at times $t_3 = rMQLT$, where $0 \leq r \leq R-1$ and the temporal frame-acquisition interval $T_s = QMLT$. For 3-D imaging, r corresponds to a spatial index ΔX_3, but for strain imaging a fixed-object region is scanned ($\Delta X_3 = 0$). The indices are nested as follows: $g(\mathbf{t}) = g(t_1, t_0, t_2, t_3) = g(\ell T, qLT, mQLT, rMQLT)$, and the acquisition time is $t' = t_1 + t_0 + t_2 + t_3 = (\ell + (q + (m + rM)Q)L)T$. Of course, the reconstruction operator $\mathcal{O}$ and coordinate transformation matrix $\mathbf{B}$ will change depending on the technique, but $\mathcal{O}$ is nonlinear for all the techniques described above. With the above data structure, it is easy to allow the object function to change with time. This situation occurs with blood flow and tissue deformation. The object function assumes the form $\gamma(\mathbf{x}, t'[\ell, q, m, r])$, therefore,

$$g[\ell, q, m, r] = \int dx h(\mathbf{x}, t')\gamma(\mathbf{x}, t') + n[\ell, q, m, r] \tag{54}$$

Re-mapping the four-dimensional matrix into a vector gives the compact form similar to (51):

$$\mathbf{g} = \mathcal{H}\{\gamma(\mathbf{x}, t')\} + \mathbf{n} \tag{55}$$

Although objects and images are naturally functions of space and time, we believe it is easier to consider echo data acquired serially strictly as a function of time. Yet,

by creating a time vector, data can be conveniently partitioned into segments that intuitively correspond to spatial and temporal coordinates of the reconstructed image.

3.10. Appendix B

Here we summarize the derivation of (27). From (26), the homogeneous wave equation in inhomogeneous media, multiply by -1 and add:

$$\frac{1}{\rho_o(\mathbf{x})}\nabla^2 p - \kappa_o \frac{\partial^2 p}{\partial t^2} \tag{56}$$

then multiply by ρ_o to find that

$$\nabla^2 p(\mathbf{x},t) - \frac{1}{c^2}\frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} = \frac{\gamma_\kappa(\mathbf{x})}{c^2}\frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} + \nabla\cdot(\gamma_p(\mathbf{x})\nabla p(\mathbf{x},t)) \tag{57}$$

where $\gamma_\kappa(\mathbf{x}) = (\kappa(\mathbf{x}) - \kappa_o)/\kappa_o$, $\gamma_p(\mathbf{x}) = (\rho(\mathbf{x}) - \rho_o)/\rho(\mathbf{x})$ and $c^2 = 1/\rho_o\kappa_o$. Using $p(\mathbf{x},t) = p_\omega(x)exp(i\omega t)$, the following time-independent solution to (56) is found:

$$\nabla^2 p_\omega(x) - k^2 p_\omega(x) = -q_\omega(x) \tag{58}$$

where

$$q_\omega(\mathbf{x}) = k^2\gamma_\kappa(\mathbf{x})p_\omega(\mathbf{x}) - \nabla\cdot(\gamma_\rho(\mathbf{x})\nabla p_\omega(\mathbf{x})) \tag{59}$$

Eq. (57) is still a homogeneous wave equation as it has no sources or sinks. The terms on the right describe scattering sources that redirect energy but do not add or subtract from the total. A solution to (57) can be found using the Greens function approach.[49] The total field is the sum of the incident and scattered fields:

$$p_\omega(\mathbf{x}) = p_{\omega_i}(\mathbf{x}) + p_{\omega_s}(\mathbf{x}) \tag{60}$$

where

$$p_{\omega_i}(\mathbf{x}) = \int_V d\mathbf{x} q_\omega(\mathbf{x}) G(\mathbf{r}|\mathbf{x}) \tag{61}$$

is the scattered field. The Green's function is defined as

$$G(\mathbf{r}|\mathbf{x}) = \frac{1}{4\pi} \frac{exp(-ik|\mathbf{r}-\mathbf{x}|)}{|\mathbf{r}-\mathbf{x}|} \tag{62}$$

From (58) and (60) we find the scattered field as:

$$\begin{aligned} p_{\omega_s}(\mathbf{r}) &= \int_V d\mathbf{x} \left\{ k^2 \gamma_\kappa(\mathbf{x}) p_\omega(\mathbf{x}) G_\omega(\mathbf{r}|\mathbf{x}) - G_\omega(\mathbf{r}|\mathbf{x}) \nabla \cdot (\gamma_\rho(\mathbf{x}) \nabla p_\omega(\mathbf{x})) \right\} \\ &= \int_V d\mathbf{x} \left\{ k^2 \gamma_\kappa(\mathbf{x}) p_\omega(\mathbf{x}) G_\omega(\mathbf{r}|\mathbf{x}) + \gamma_\rho(\mathbf{x}) \left[\nabla p_\omega(\mathbf{x}) \cdot \nabla G_\omega(\mathbf{r}|\mathbf{x}) \right] \right\} \end{aligned} \tag{63}$$

The last expression was obtained by using the product rule for differentiation followed by Gausss theorem. This is the scattered pressure field at points on the receiving transducer surface. It is a function of the total field $p_\omega(x)$, which is well approximated by the incident field $p_{\omega_i}(x)$ under a weak-scattering hypothesis as is reasonable in biological tissue. From[48] the incident field from a quasi-planar surface is given as:

$$p_i(x,t) = i\omega_o ckU(\omega) exp(i\omega t) A(\mathbf{x},t) \tag{64}$$

where A is the velocity potential described in (28). Integrating the scattered pressure (62) over the transducer surface S, and applying the weak scattering approximation and (63), the expression (27) for the force exerted on the transducer is obtained.

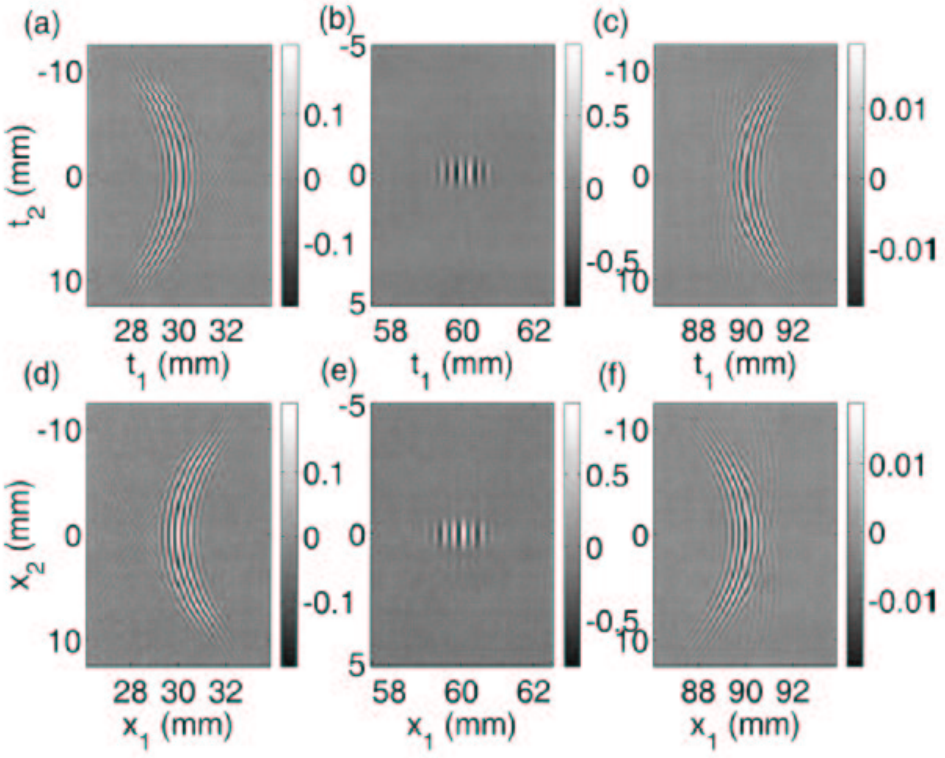

Figure 2. Point-spread (a)(c) and spatial sensitivity (d)(e) functions for a 3 MHz array transducer. Images of psfs represent spread from one scatterer positioned at 30 (near field), 60 (focal region), and 90 mm (farfield), respectively. Similarly, spatial sensitivity functions are shown for pulse-echo temporal intervals corresponding to times labeled by half the pulse-echo path distance: 30, 60, and 90 mm, respectively. The psf and ssf functions were normalized by the maximal focal region magnitude. Also, the axes for the psf images were plotted in millimeters for convenience in comparing with the ssfs. The array consisted of 64 active elements of width one wavelength, height 5 mm, and inter-element gap separation of 0.1 mm. The electronic focus of the array was 60 mm. The excitation was taken as a four cycle pulse filtered by the aperture electromechanical response simulated as a four-cycle Hanning-windowed pulse. The psf images were generated with Field II, then filtered to include the effects of attenuation and scattering. The medium was taken to have sound speed of 1540 m/s, and attenuation coefficient of $0.3\text{dB/cm/MHz}^{1.1}$. The ssfs were generated from an ensemble of psfs as described above.

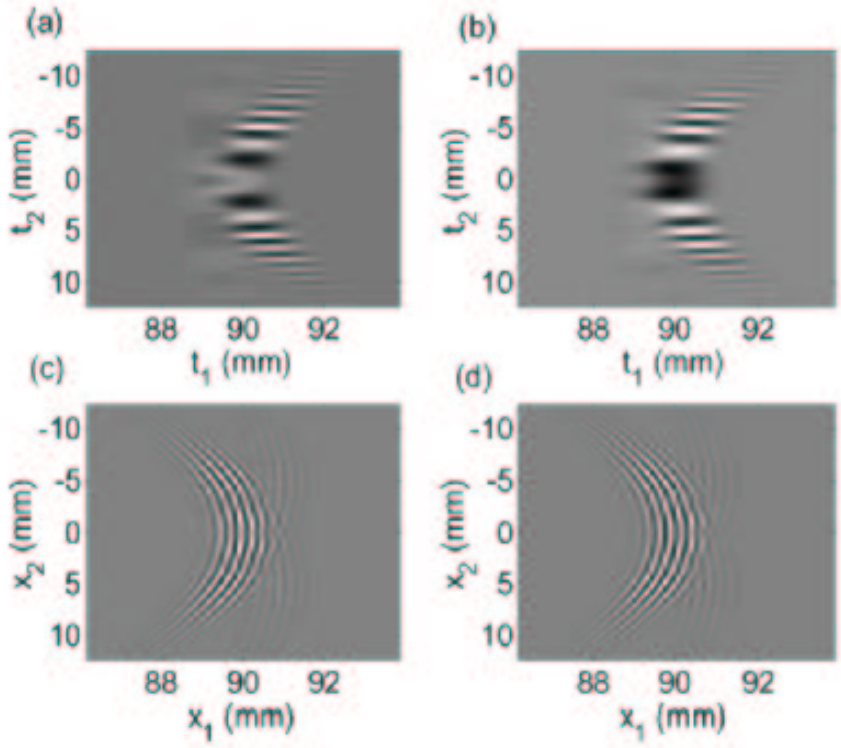

Figure 3. Normalized in-phase (a) and quadrature (b) point-spread functions, and in-phase (c) and quadrature (d) spatial sensitivity (c)(d) functions for the array transducer described in Fig. 2. Images of I and Q psfs represent spread from a scatterer at a distance of 90 mm (farfield). Similarly, I and Q spatial sensitivity functions are shown for pulse-echo temporal intervals corresponding to 90 mm, which is half the pulse-echo path distance.

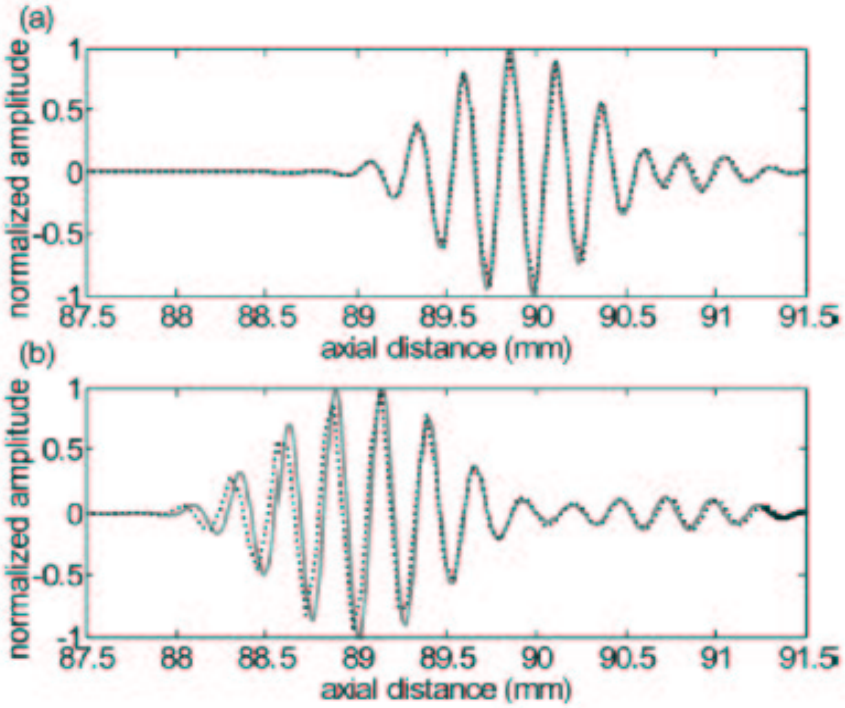

Figure 4. Comparisons between the normalized farfield ssf and time-reversed psf A-scan lines (a) on axis and (b) 7.6 mm off axis show that the quality of approximation of (41) depends on experimental parameters including position. The parameters were the same as in Fig. 3.

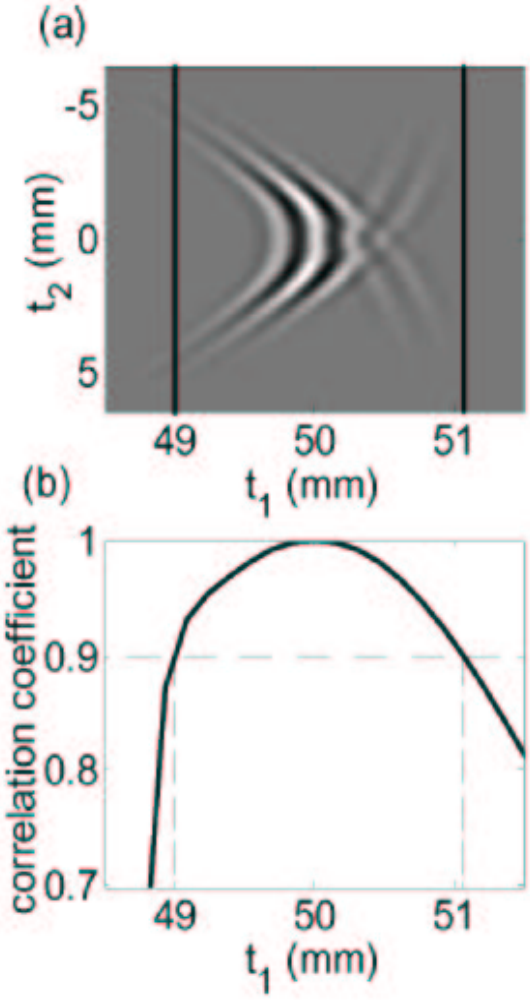

Figure 5. A method for determining the size of an isoplanetic patch. The vertical blank lines in (a) represent the point at which the correlation coefficient as defined by (20) reach 90% threshold. In this example, because most of the energy is contained within this region, the local LSIV approximation may be expected to be a good one. Point target was at 50 mm, and focal region was 60 mm. A two-cycle excitation was used, and the apertures electromechanical impulse response was simulated as a two-cycle Hanning-windowed pulse. Otherwise, simulation parameters were the same as in Fig. 2.

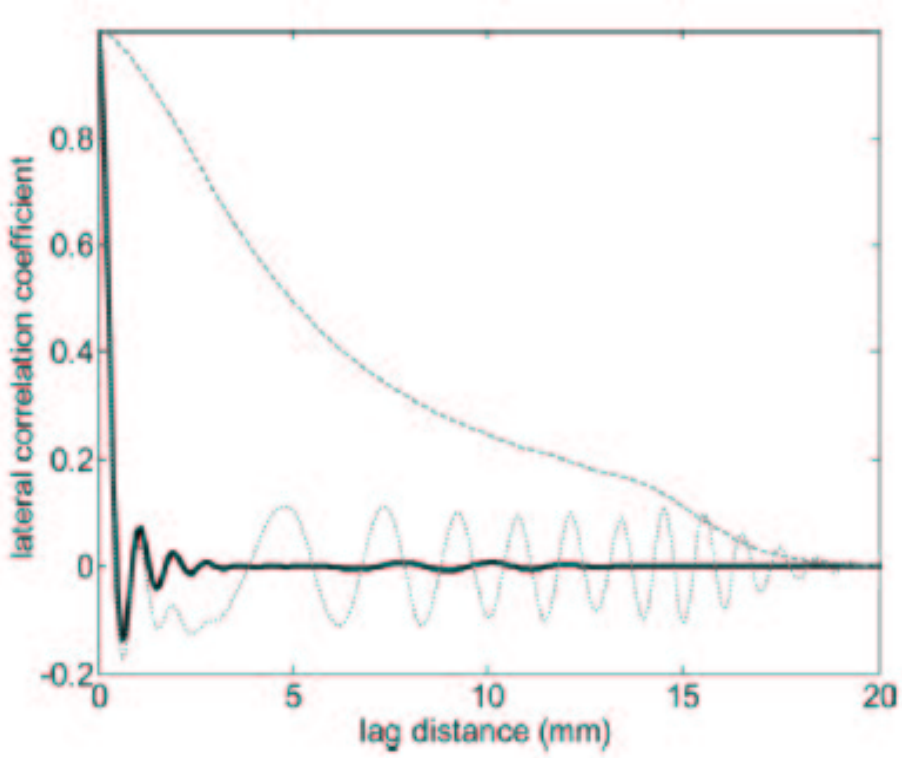

Figure 6. Predicted correct and erroneous correlation coefficients of in-phase image data from lateral scan lines at 30 mm axial depth for a transducer with focus at 60 mm. Simulation parameters were otherwise the same as in Fig. 2. Solid line: the predicted true correlation coefficient using the spatial sensitivity functions. Dashed line: an erroneous predicted lateral correlation coefficient arrived at by ignoring phase oscillations. Dotted line: the erroneous lateral correlation coefficient arrived at by using the in-phase point-spread functions rather than in-phase sensitivity functions. The true nearfield speckle is predicted to be very narrow in the lateral direction due to wavefront curvature. Not shown is the in-phase correlation coefficient predicted by a phase-shifted, time-reversed RF psf. It differed from the correct lateral correlation coefficient by less than 1%.

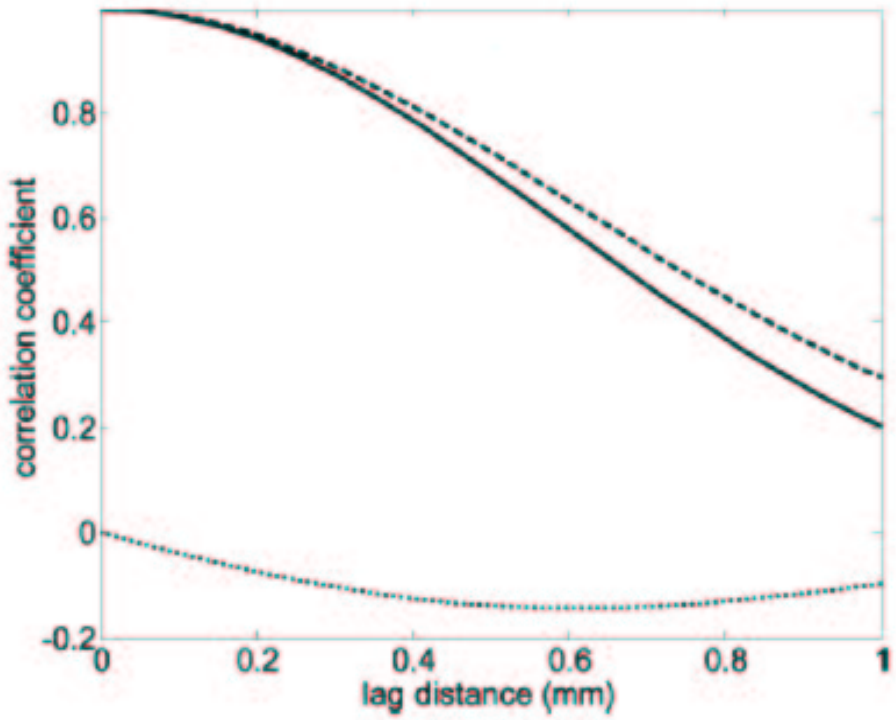

Figure 7. Normalized covariances $K_{II}(t, t+\Delta t)$, solid and dashed lines, and $K_{IQ}(t, t+\Delta t)$, dotted line, of in-phase and quadrature signals along axial scan lines as a function of lag distance $\Delta x = c\Delta t/2$. Calculations were based on 2-D field distributions from a focused array transducer with focus at 6 cm, and $ct/2 = 90$ mm. Otherwise, parameters were the same as in Fig. 2. The dashed line represents the computation using the isoplanetic approximation (modeling I and Q ssfs by time-reversed, phase-shifted psfs). The solid line is the computation with the full shift-variant theory. Note from the dotted line that in-phase and quadrature correlations exist. This may be explained as coupling between direct and edge waves, as shown in Fig. 33.

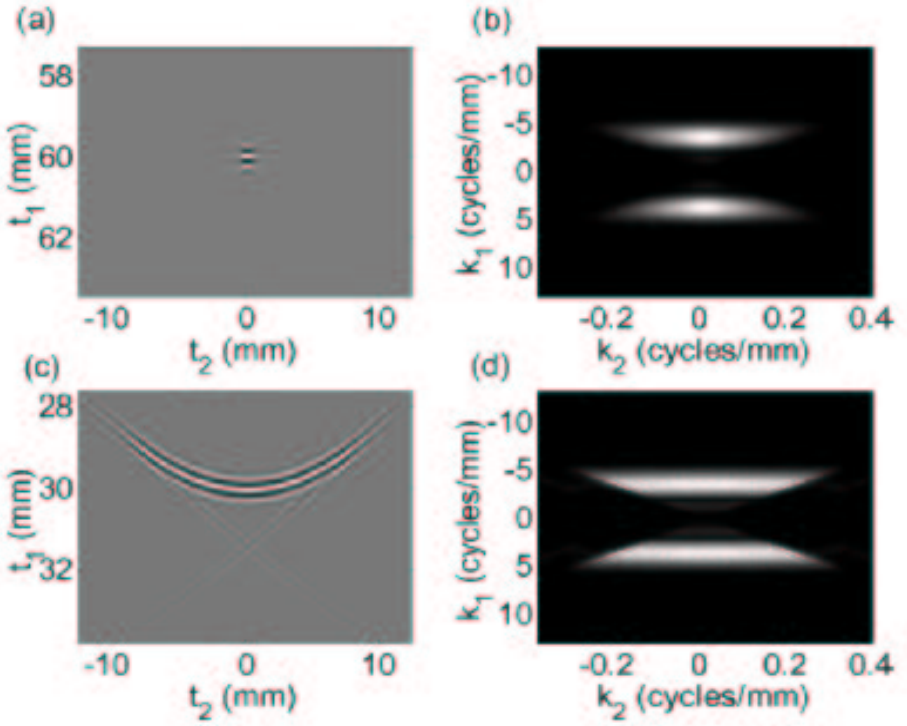

Figure 8. (a) Focal and (c) nearfield RF psfs due to a 3 MHz fixed focus linear array of height 0.5 cm, with 64 active elements of width λ separated by distances of 0.1 mm. Azimuthal focus was 6 cm. The beam was electronically swept laterally across the point target located at (a) 6 cm and (c) 3 cm. (b) and (d) are the k-space representations of (a) and (c), respectively. A two-cycle excitation was used, and the apertures electromechanical impulse response was simulated as a two-cycle Hanning-windowed pulse.

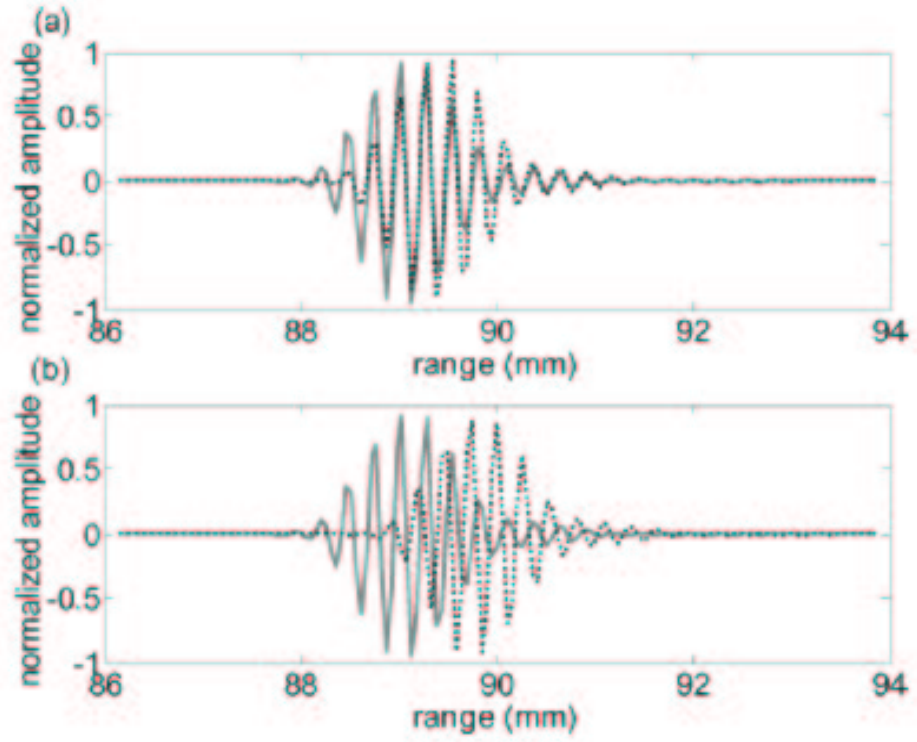

Figure 9. (a) Axial slices through in-phase ssfs. This is to illustrate that in-phase ssfs are always in phase with each other, even though the envelope may shift. (b) In-phase (solid line) and quadrature (dotted line) ssfs are not 90^o out of phase. At about the 90-mm point, the direct wave of the quadrature ssf is almost 180^o out of phase with the edge wave of the in-phase ssf.

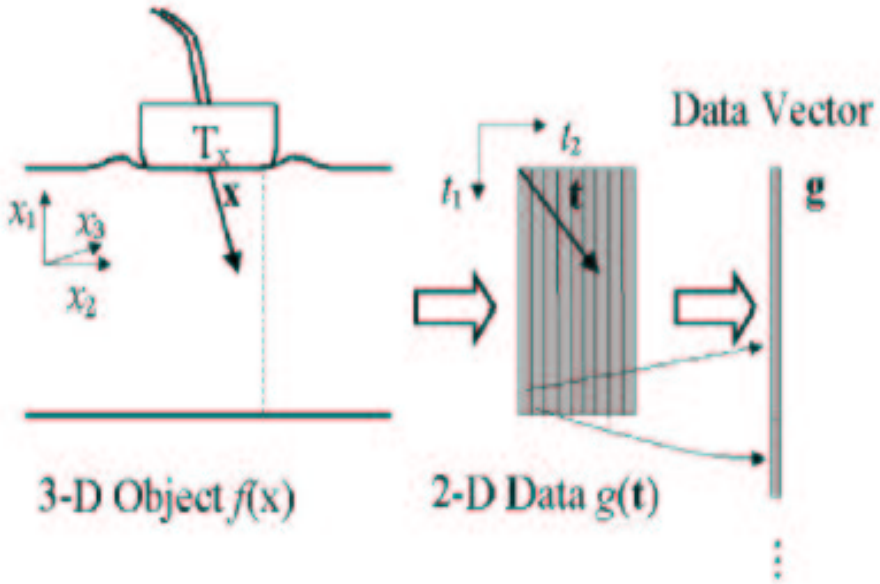

Figure 10. Illustration of geometries for the object and data vectors using a linear array transducer.

4. CHALLENGES IN ULTRASONIC DETECTION THEORY

So far we have spent considerable time setting up the linear systems framework that we will need for an analytical treatment of ultrasound detection theory. We now begin work on the ideal observer for ultrasonic detection tasks, and describe some of the challenges in the analysis.

4.1. Task Definition

It is the statistical properties of the echo signal that give us information about the object imaged. The probability density functions for signal present and signal absent RF echo signals can be modeled as zero-mean multivariate normal densities

$$pdf(\mathbf{g}|+) = \frac{exp(-\frac{1}{2}\mathbf{g}^t\mathbf{K}_+^{-1}\mathbf{g})}{(2\pi)^{N/2}\sqrt{det\mathbf{K}_+}} \tag{65}$$

and

$$pdf(\mathbf{g}|-) = \frac{exp(-\frac{1}{2}\mathbf{g}^t\mathbf{K}_-^{-1}\mathbf{g})}{(2\pi)^{N/2}\sqrt{det\mathbf{K}_-}} \tag{66}$$

Here the covariance matrices $\mathbf{K}_+ = \langle \mathbf{g}\mathbf{g}^t|+\rangle$ and $\mathbf{K}_- = \langle \mathbf{g}\mathbf{g}^t|-\rangle$ describe the image textures for the signal present and signal absent hypotheses respectively. A lesion present, thus, for example, may have a region of increased variance relative to the surrounding tissue. The echo data vector g is assumed to be a $NM \times 1$ column vector, that is formed by rastorizing M A-scan echos of length N.

For the detection task that we are interested in, the signal is known statistically - as characterized by the covariance $\Delta\mathbf{K} = \mathbf{K}_+ - \mathbf{K}_-$. The location of the lesion

is assumed to be known exactly, as is its shape and contrast. The background is characterized statistically as described by the covariance $\mathbf{K}_-$.

4.2. Ideal Observers of RF Echo Signals

The ideal observer makes its decisions based on thresholding the optimal decision statistic, the log-likelihood ratio is given as

$$\begin{aligned}\lambda_{Log-likelihood}(\mathbf{g}) &= log\left(\frac{pdf(\mathbf{g}|+)}{pdf(\mathbf{g}|-)}\right) \\ &= \frac{1}{2}\mathbf{g}^t\left(\mathbf{K}_+^{-1}-\mathbf{K}_-^{-1}\right)\mathbf{g}+\frac{1}{2}log\left[det\left(\mathbf{K}_+\mathbf{K}_-^{-1}\right)\right]\end{aligned}$$

We can ignore the second term as it is a data-independent constant. The ideal observer test statistic gives the strategy for optimal detection. The strategy is to compute the difference between the squared magnitudes of the pre-whitened data. Later on we comment on the importance of the ideal observer strategy for improving human observer performance by image processing. Eq. 67 represents a signal known statistically (SKS) quadratic task. By quadratic task we mean the test statistic is a quadratic form of the data vector $\mathbf{g}$. At this point we can compute the SNR of the test statistic using 1. To do so we need the moments of the test statistic.

$$SNR_\lambda^2 = \frac{\left\{tr\left[(\mathbf{K}_+^{-1}-\mathbf{K}_-^{-1})(\mathbf{K}_+-\mathbf{K}_-)\right]\right\}^2}{2\left(tr\left\{\left[(\mathbf{K}_+^{-1}-\mathbf{K}_-^{-1})\mathbf{K}_+\right]^2\right\}+tr\left\{\left[(\mathbf{K}_+^{-1}-\mathbf{K}_-^{-1})\mathbf{K}_-\right]^2\right\}\right)} \tag{67}$$

where $tr[\cdot]$ is the trace of the matrix. The expression in the denominator made use of Isserlis's formula for fourth order moments of Gaussian distributed signals.[65] Eq. 10 describes how echo signal textures due to the system and object inter-relate to affect task performance. It is important because it extends the Smith-Wagner

theory[29] to include pixel correlations, and in fact the SNR reduces to the Smith-Wagner model when covariances are diagonal and stationary, and pixels are sampled on the same scale as correlation lengths. It is not a very intuitive model, however, and could be computationally expensive to evaluate due to the inverse covariances needed. Our ultimate goal is to gain more intuition of ideal observer performance and so we consider how other approaches could complement this analysis.

4.3. Clarkson-Barret Approach

Here we summarize an approach to SNR analysis taken by Clarkson and Barrett.[36] One contribution of our work is to point out the applicability of the Clarkson-Barrett theory to ultrasound performance assessment. We shall also use this theory later on to understand a more intuitive picture of image quality.

The ideal observer is dependent on the probability distributions of the test statistic - conditioned on the hypothesis that the signal is present or absent. Alternatively, instead of *pdf*s, moment-generating functions[13] for the log-likelihood may be used. They are essentially the Laplace transform of the *pdf*s,

$$M_+(\beta) = \langle exp(\beta\lambda)\rangle_+$$

$$M_-(\beta) = \langle exp(\beta\lambda)\rangle_+$$

and hence are also related to the characteristic functions (Fourier transform) of the *pdf*s (likelihoods). Because probability must be conserved (probability of lesion present + probability of lesion absent = 1), all the information about the area under

the ROC curve is contained in one of the moment-generating functions. The moment-generating functions can consequently be written in terms of a single function, called the likelihood generating function $G(\beta)$:

$$\begin{aligned} M_+(\beta) &= exp\left[\beta(\beta-1)G(\beta-1/2)\right] \\ M_-(\beta) &= exp\left[\beta(\beta+1)G(\beta+1/2)\right] \end{aligned}$$

Both moment-generating functions and likelihood generating functions are useful for quantifying detectability of the signal with stochastic calculations. Different kinds of metrics can be used to quantify detectability. We have already discussed SNR_λ. The area under the ROC curve (AUC) is also a useful measure of detectability. Clarkson and Barrett show that SNR_λ has the unusual behavior that it does not increase without bound as signal strength increases. This SNR is also not invariant to monotonic transformation of the decision variable, whereas the AUC is. Another SNR is given by

$$SNR_{G(0)} \equiv \sqrt{2G(0)} \tag{68}$$

where

$$G(0) = -4logM(1/2) \tag{69}$$

This SNR is invariant under monotonic transformations of the decision variable, and is related to the Bhattacharyya distance, a metric quantifying the "distance" between two general probability density functions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$:

$$d_B(p_1,p_2) = -log\left\{\int d\mathbf{x}[p_1(\mathbf{x})p_2(\mathbf{x})]^{1/2}\right\} \tag{70}$$

When the two distributions have no overlap, there is very good separability. In this case, the integral tends to zero, and the Bhattacharyya distance becomes infinite. When there is complete overlap, the distributions are identical and normalization constrains the integral to 1. Consequently the Bhattacharyya distance tends to zero. For our task, the essential moment generating function is

$$M_{-}(\beta) = \frac{\sqrt{det\mathbf{K}_{-}}^{\beta-1}\sqrt{det\mathbf{K}_{+}}^{-\beta}}{\sqrt{det\left[(1-\beta)\mathbf{K}_{-}^{-1}+\beta\mathbf{K}_{+}^{-1}\right]}} \tag{71}$$

and the corresponding signal-to-noise ratio is

$$SNR^{2}_{G(0)} = 2log\left\{\frac{\left[det(\mathbf{K}_{+}+\mathbf{K}_{-})\right]^{2}}{2^{2N}det\mathbf{K}_{+}det\mathbf{K}_{-}}\right\} \tag{72}$$

Intriguingly and conveniently this SNR does not require inversion of any covariance matrices. It's computational evaluation may nevertheless be prohibitive for non-stationary covariance matrices due to expensive determinant operations.

4.4. The Need for an Analytically or Computationally Tractable Performance Theory

Neither the test statistic 67 nor the expressions for ideal observer SNR (Eqs. 67 and 72) are easily computable due to the enormous sizes of the covariance matrices. For echo data of dimension 128×128, the data vector is a column vector of size 16384×1, and the covariance is thus 16384×16384! In expressions for the test statistic 67 and ideal observer SNR 67, inversion of the signal present and signal absent covariance matrices is required. Inversion of $\mathbf{K}_{-}$ is not a problem when the

the system is linear shift-invariant since the we assume the background is a stationary random process. This means that $\mathbf{K}_-$ has a block-Toeplitz structure - that is well approximated by a block-circulant structure, and hence diagonalizable by a discrete Fourier transform.[23] It should be noted that when diagonalizing, we do not actually need to use 16384×16384 matrices. Rather, we know that the eigenvalues are simply the power spectrum of the echo signal.

Inversion of $\mathbf{K}_+$ is much more complicated, since the statistical properties of the signal present image are position-dependent - and thus not stationary. This means that conventional DFT operations cannot directly be used to diagonalize $\mathbf{K}_+$. Trying to invert $\mathbf{K}_+$ with standard matrix inversion techniques are impractical due to computational complexity and instability. Regarding computational complexity, for example, to solve a linear system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is $N \times N$ using LU factorization requires $\mathcal{O}(N^3)$ operations - thus the total cost for matrix inversion is of order $\mathcal{O}(N^4)$.[66] Alternatively, eignevalue decomposition methods such as Hessenberg QR Iteration has complexity on the order of $\mathcal{O}(N^3)$[66].[67] Stability is better for large condition numbers. The condition number of a matrix $\mathbf{A}$ can be defined as $\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$ where $\|\cdot\|$ represents the L2 norm. An alternative definition is the ratio of the largest and smallest eigenvalues. There are a number of algorithms available for estimating condition number. $\mathbf{A}$ is ill-conditioned if $\kappa(A) \geq 1$ and well-conditioned otherwise.

The Clarkson-Barrett expression for $SNR_{G(0)}$ 72 does not require inversion of signal present covariance matrices, but does require computation of determinants. With eigenvalue techniques such as Hessenberg QR Iteration, the computational complex-

ity is of order $\mathcal{O}(N^3)$ - again a challenging task when matrices are 16384×16384. Stability is also equally challenging.

Approaches to addressing these challenges will be discussed in the following chapters in this dissertation.

5. GENERALIZED NEQ FOR ASSESSMENT OF ULTRASOUND IMAGE QUALITY

5.1. Introduction

5.2. Additive Hotelling Ideal Observer

One unsatisfying aspect of the derived figures of merit for the quadratic task of detection is that the imaging system parameters are buried inside covariance matrices. Hence interpretation and intuition for system optimization is not obvious. Moreover, for the problems we are interested in, the object statistics are non-stationary (statistical properties of the image are location-dependent), hence the echo signals themselves are non-stationary. The advantage of stationarity is that diagonalization is natural and algorithmically efficient with a discrete Karhunen-Loeve expansion.[23] Lacking such stationarity, computational evaluation of image quality metrics can be prohibitive.

Another approach to modeling the signal detection problem is as follows. Consider an object function for the signal absent that is a zero-mean stationary stochastic process $\gamma_b(x)$ with covariance $\mathbf{K}_b$. For the signal present case consider that the object function can be modeled as the random background $\gamma_b(x)$ plus an additive signal $\Delta\gamma_b(x)$ that is one deterministic realization of a random process (the statistical properties of this random process may vary from point to point in space). In essence, we are using the concept that a multiplicative signal in the variance can be represented

as an additive signal. The advantage of this is that we are now back to a problem that is well known in the image quality literature: an additive signal $\mathbf{s} = \mathbf{H}\Delta\gamma_b(x)$ and stationary covariance $\mathbf{K} = \mathbf{H}^t\mathbf{K}_b\mathbf{H} + \mathbf{K}_n$, where $\mathbf{K}_n$ is the noise covariance. The problem is descriptive of hyper-echoic lesions (lesions where the backscattered intensity is greater than the surrounding tissue background). The optimum linear observer is the Hotelling observer[11] which has test statistic

$$\lambda_{Hot} = \mathbf{s}^t\mathbf{K}^{-1}\mathbf{g} \tag{73}$$

with SNR given by

$$SNR^2_{Hot} = tr\left[\mathbf{K}^{-1}\mathbf{s}\mathbf{s}^t\right] \tag{74}$$

It has been shown that this SNR is identical to the ideal observer SNR for non-random signals in normally distributed random backgrounds. The strategy for the Hotelling observer is to perform matched filtering with a pre- whitening step. Assuming local shift-invariance, $\mathbf{K}$ is stationary, and a frequency space description of the SNR is possible via a Fourier transform. Additionally, we average over signal realizations to obtain:

$$SNR^2_{Hot} = \int d\mathbf{u} \frac{\langle|\Delta\Gamma(\mathbf{u})|^2\rangle_\gamma |H(\mathbf{u})|^2}{|H(\mathbf{u})|^2 S_{\gamma_b}(\mathbf{u}) + S_n(\mathbf{u})} \tag{75}$$

Similar averaging of SNR^2 has been done before[30] except over location uncertainty. Here $\Delta\Gamma(\mathbf{u})$ is the Fourier description of the object function signal $\Delta\gamma(\mathbf{x})$, $S_{\gamma_b}(\mathbf{u})$ is the power spectrum of the background object function process, and $S_n(\mathbf{u})$ is the noise power spectrum. $H(\mathbf{u})$ is the system response function. The integration is

done out to the extent of the sampling frequency. This is an ensemble average over a generalized Wagner-Brown model of detectability[6830].[27]

Before examining the interesting properties of this SNR, we first examine the expectation value of the magnitude of the Fourier domain target signal in the numerator of the integral. Consider that the spatial domain target signal is given by $\Delta\gamma(\mathbf{x}) = w(\mathbf{x})\xi(\mathbf{x})$, where w is a deterministic window function and ξ is stochastic process. The purpose of writing $\Delta\gamma$ in this form is to characterize the location dependence of target signal statistical properties. For example, w could be a disk signal that is unity inside the lesion and zero outside the lesion. In general, the Fourier domain stochastic average over target signal variations can be written as

$$\begin{aligned} S_{\Delta\Gamma}(\mathbf{u}) &= \frac{1}{X}\langle|\Delta\Gamma(\mathbf{u})|^2\rangle = \frac{1}{X}\langle|W(\mathbf{u}) * \Xi(\mathbf{u})|^2\rangle \\ &= \frac{1}{X}\int\int d\mathbf{p}d\mathbf{q}W(\mathbf{p})W^*(\mathbf{q})\langle\Xi(\mathbf{u}-\mathbf{p})\Xi(\mathbf{u}-\mathbf{q})\rangle\rangle \end{aligned} \tag{76}$$

where W and Ξ are the Fourier transforms of w and ξ respectively, and Ξ is the spatial region of support over the image (the image area or volume).

At this point, we look at some simplifying assumptions to compare our results with those of the Smith-Wagner theory. If in particular ξ is a white Gaussian noise (WGN) process with variance σ_ξ^2, then Ξ is also WGN. The WGN assumption is a good one when spatial variations of density and compressibility are on a much smaller scale than the resolution of the imaging system and smaller than the sampling intervals used. With this assumption,

$$S_{\Delta\Gamma}(\mathbf{u}) = \sigma_\xi^2 \int d\mathbf{x}|w(\mathbf{x})|^2 \tag{77}$$

by Parseval's theorem. If, for example, we model disk signals as Smith and Wagner did, such that $w(x)$ is unity where the lesion is present and zero where it is absent, we have $S_{\Delta\Gamma}(\mathbf{u}) = \sigma_\xi^2 A$, where A is the area of the lesion (or volume of the lesion if considering 3-dimensions). Thus, when the object statistics can be treated as WGN, is a constant, that is proportional to the object variance and the area of the lesion. Hence,

$$SNR^2_{GNEQ} \equiv \langle SNR_I \rangle_\gamma = \sigma_\xi^2 AX \int d\mathbf{u} \frac{|H(\mathbf{u})|^2}{|H(\mathbf{u})|^2 S_b(\mathbf{u}) + S_n(\mathbf{u})} = \sigma_\xi^2 AX \int d\mathbf{u} GNEQ(\mathbf{u}) \tag{78}$$

We have written this quantity as an integration over what has been termed the Generalized Noise-Equivalent Quanta[68] or $GNEQ$, defined as the integrand of the third term. Noise-Equivalent Quanta (NEQ) has historical origins with Shaw[24] and others, as well as the Wagner-Brown theory of detectability.[27] For photon imaging systems NEQ represents the frequency-specific density of quanta at the input of an ideal detection system that would yield the same output noise as the real system under evaluation. Generalized NEQ as described by Barrett and colleagues[68] provides provision for a stochastic background texture of the object. For ultrasound systems the $GNEQ$ quantity is a measure of the spatial frequency sensitivity of detecting a signal in a background texture and in the presence of electronic noise. For photon imaging modalities, the ideal observer detectability is given by an integral over the frequency-domain of NEQ (or $GNEQ$) times the squared magnitude of the Fourier transform of the signal. For ultrasound systems, the $GNEQ$ is not weighted by the magnitude of the Fourier transform of the signal shape, but rather by the spectral

variance of the target signal. This is to be expected since ultrasound systems are not sensitive to the magnitude of density or compressibility but rather depend on the variance (spatial fluctuations) of these quantities. The entire $GNEQ$ spectrum can be used to characterize ultrasound systems for a particular tissue type in a target-independent manner. Defining as the integrated background power spectral density, we can also re-write the SNR as:

$$SNR^2_{GNEQ} = C_{obj} A \int d\mathbf{u} \frac{MTF^2(\mathbf{u})}{MTF^2(\mathbf{u}) \frac{S_{\gamma_b}(\mathbf{u})}{X\sigma_b^2} + \frac{1}{S/N(\mathbf{u})}} \tag{79}$$

where C_{obj} is the object function contrast, defined as $C_{obj} = \sigma_\xi^2/\sigma_b^2$ and MTF is the modulation transfer function defined as

$$MTF(\mathbf{u}) = |H(\mathbf{u})|/|H|_{max} \tag{80}$$

where $|H|_{max}$ is the maximum value of $|H(\mathbf{u})|$. The background signal-to-noise ratio is defined as:

$$S/N(\mathbf{u}) = |H|^2_{max} X\sigma_b^2/S_n(\mathbf{u}) \tag{81}$$

and is a measure of the backscattered signal strength of the object medium relative to the noise level. Note the connections between our expression for SNR and the Smith-Wagner theory: both SNR^2 measures are proportional to lesion area. Curiously, $[SNR_{GNEQ}]^2$ is proportional to contrast, yet in the Smith-Wagner theory, the SNR^2_λ is proportional to the square of object contrast. Both metrics, however, reveal that task performance is improved with improved contrast.

One elegant feature of this formalism is the direct dependence of system and noise properties on the detectability. Also, the description is in the Fourier domain,

a natural choice for analyzing ultrasound imaging systems because focused fields in the Fresnel region and unfocused apertures in the far-field (Fraunhoffer) region have k-space descriptions equivalent to the Fourier transform of the aperture. Also, the quantities involved in the $GNEQ(\mathbf{u})$ are all measurable quantities, so this could provide a convenient way to characterize a system design experimentally! To do so we would need to measure the $MTF(\mathbf{u})$ and the normalized background plus noise power spectrum ($NBNPS$), which we define as:

$$NBNPS(\mathbf{u}) = [|H(\mathbf{u})|^2 S_b(\mathbf{u}) + S_n(\mathbf{u}]/|H|^2_{max} \tag{82}$$

then take the ratio of the two:

$$GNEQ(\mathbf{u}) = MTF^2(\mathbf{u})/NBNPS(\mathbf{u}) \tag{83}$$

Wire or bead phantoms can be used to estimate the point-spread functions - and consequently the MTF. It may be necessary to acquire a small number of images to average over noise properties. Assuming local ergodicity, a patch of one or a few (RF) images may be used to estimate the background+noise power spectrum for a particular field region. Once we know these properties, the integrated $GNEQ$ or $IGNEQ$ as we shall call it may be computed:

$$IGNEQ = \int d\mathbf{u} GNEQ(\mathbf{u}) \tag{84}$$

This metric may apply to media where object statistics can be treated as nearly white, such as standard phantoms and many types of tissues. Conveniently, this metric is independent of the signal size and contrast, and characterizes the performance of

the imaging system with a single number. It may be normalized by to obtain a number useful for simulation studies, which we shall call the normalized $IGNEQ$. Tissues that have complex striations or variable patches of brightness may make target discrimination more difficult. Object texture can be accounted for in the theory and measured experimentally.

5.3. Monotonicity with the Full Quadratic Task

Recall that when we modeled both the lesion and its surroundings as Gaussian stochastic processes the resulting ideal observer test statistic was a quadratic form in the data. In contrast, when we modeled the lesion as an additive deterministic signal, the test statistic was linear in the data. We may very well expect that performance metrics for these two detection tasks may not exhibit identical properties. Consequently, we now examine whether the figures of merit derived for both tasks give similar information about diagnostic performance. To do so we choose to examine the trends of both the Clarkson-Barrett SNR and the $GNEQ$ theory with varying system and object parameters. Consider the case where both $\mathbf{K}_+$ and $\mathbf{K}_-$ are simultaneously diagonalizable with a Karhunen-Loeve transformation. In assuming this we consider the case where the size of the lesion is the same size as the window [$w(\mathbf{x}) = 1$ everywhere], and also that the system is locally shift-invariant. The result is a spectral description:

$$SNR^2_{G(0)} = 2log\left\{\frac{\prod_{m=1}^{M}\left[|H(\mathbf{u}_m)|^2(S_+(\mathbf{u}_m)+S_-(\mathbf{u}_m))+2S_n(\mathbf{u}_m)\right]^2}{2^{2M}\prod_{m=1}^{M}\left[|H(\mathbf{u}_m)|^2S_+(\mathbf{u}_m)+S_n(\mathbf{u}_m)\right]\prod_{m=1}^{M}\left[|H(\mathbf{u}_m)|^2S_-(\mathbf{u}_m)+S_n(\mathbf{u}_m)\right]}\right\} \tag{85}$$

where $S_+(\mathbf{u})$ and $S_-(\mathbf{u})$ are the eigenvalues (spectrum) of $\mathbf{K}_+$ and $\mathbf{K}_-$ respectively, and $\mathbf{u}$ is a (spatial) frequency vector of length M. Note that $S_-(\mathbf{u}) = S_b(\mathbf{u})$ and $S_-(\mathbf{u}) = S_b(\mathbf{u}) + S_{\Delta_\gamma(\mathbf{u})}$, where $S_b(\mathbf{u})$ is the spectrum of the background object function, and $S_{\Delta_\gamma(\mathbf{u})}$ is the spectrum of the target. Also, $S_n(\mathbf{u})$ is the noise-power spectrum (NPS). The spectral description now gives a means for comparing with SNR_{GNEQ}. In particular, we shall examine trends of both $SNR_{G(0)}$ and SNR_{GNEQ} for parameters of bandwidth, noise power, and target signal power.

To investigate the behavior of the SNR with system bandwidth (BW), consider for simplicity an ideal low-pass system response:

$$H(\mathbf{u}) = \begin{cases} 1 & \mathbf{u} \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{86}$$

where Ω is the set of values of spatial frequency that determines the bandwidth. Suppose there are M_Ω spatial frequency elements in Ω. Suppose also that object and noise statistics are white: $S_+(\mathbf{u}) = S_+$, $S_-(\mathbf{u}) = S_-$, and $S_n(\mathbf{u}) = S_n$ for all $\mathbf{u}$. Then

$$SNR^2_{G(0)} = 2M_\Omega log\left\{\frac{[S_+ + S_- + 2S_n]^2}{2^{2M/M_\Omega}\,[S_+ + S_n]\,[S_- + S_n]}\right\} \tag{87}$$

Consequently, the SNR_2 is proportional to M_Ω, which is simply proportional to $|\Omega|$, the system bandwidth.

To show that $SNR_{G(0)}$ decreases with increasing noise, we took the derivative with respect to the noise power and showed that the result was a strictly negative quantity, indicating that the slope of the graph of $SNR_{G(0)}$ with respect to noise was negative (calculation not shown). As the noise grows larger without bound, both

$SNR_{G(0)}$ and SNR_{GNEQ} both tend to zero. As noise power tends to zero, $SNR_{G(0)}$ and SNR_{GNEQ} increase to a limit that is only bound by the density of sampling, and the Nyquist limit of H. If H is non-zero for frequency channels $u \in \Omega$, then as noise tends to zero:

$$SNR^2_{GNEQ} \longrightarrow_{noise \to 0} \propto \int d\mathbf{u} \frac{1}{S_b(\mathbf{u})} \tag{88}$$

As the sampling frequency grows, so does the SNR as long as H has even a small amount of signal. The same behavior is seen with $SNR_{G(0)}$:

$$SNR^2_{G(0)} \longrightarrow_{noise \to 0} \sum_{m \ni \mathbf{u}_m \in \Omega} 2log \left\{ \frac{[S_+ + S_-]^2}{2^{2M/M_\Omega} [S_+ S_-]} \right\} \tag{89}$$

As the number of frequency channels in W increases, this quantity will in general increase until both target and background object power spectra tend to the same number, for example, as they die off to zero. In the Smith-Wagner theory, noise-power could be totally neglected, and the detectability was bounded not by the sampling density, but by the number of independent speckle spots. By using the RF signal instead of the envelope signal, the ideal observer has the potential to compensate for the blurring kernel of the imaging system. Noise prevents the ideal observer from doing this completely. One interpretation of the ideal observer strategy is that it deconvolves the effects of the imaging system before performing template matching. Deconvolution cannot necessarily recover precise information lost in the null-space but it can restore image statistics useful for the detection task. An interesting observation related to deconvolution is that the argument of the SNR_{GNEQ} integral looks very much like a Weiner filter. In practice, there will always be noise, and the ideal

observer detectability will be bounded. The imaging system should be designed so as to sample at minimum at the Nyquist rate, determined by the system's MTF. Sampling more densely than the Nyquist rate will add noise power to the images (which may degrade task performance for human observers). For ideal observers, sampling at greater than the Nyquist limit will not degrade performance, but will not provide any benefit either. This is because the ideal observer will know to truncate spectral noise where there is no signal.

Lastly, we note the increasing trends of both $SNR_{G(0)}$ and SNR_{GNEQ} with increasing signal strength. SNR_{GNEQ} is proportional to target signal variance. For $SNR_{G(0)}$ the relationship with target strength is more subtle, however, by differentiating with respect to signal strength, one can show that the relevant slope is positive (calculation not shown).

The conclusion of the previous analysis is that SNR_{GNEQ}, a figure of merit for a slightly different task than the full quadratic task modeled in sections II, IV, and V gives similar information as do metrics for the quadratic task. The benefits are that small targets (lesions) can be considered, noise and system parameters can be included directly and intuitively into the analysis, and the SNR can be both easily measured and predicted with a target-independent metric.

5.4. Shift-Variant Information Maps of Ultrasonic Fields

Realistic ultrasound systems are not shift-invariant. But we can assume that they are shift-invariant over some local region. By doing so, we can compute the potential

information content of imaging a lesion at each field location. We considered a 128-element linear array transducer with 64 active elements and fixed focus at 60 mm. No aperture growth, dynamic receive focusing, elevation focusing, or apodization were employed. Attenuation was neglected, and we only consider 2 dimensions of a truly 3-D ultrasound beam sensitivity. All such effects could easily be integrated into future simulations. Point spread functions (psfs) simulated using FIELD II,[53] a public domain ultrasound simulator, are shown below in Fig. 11(a). This is a B-mode image of point scatterers located at successive axial depths separated by 5 mm intervals. At each point the MTF was computed, corresponding to the normalized 2-D FFT of the RF point-spread functions. Assuming white background and noise processes, and $S/N = 1/30$ at the focus, a normalized $IGNEQ$ value was computed at each axial point, and the results plotted in Fig. 11(b).

The results are somewhat counter-intuitive. We would have expected the focal region to be optimally informative since it has the most focused resolution size. Instead we find that the nearfield has remarkably more information potentially available. To understand this unexpected result, we compared the nearfield and focal region psfs and MTFs, as shown in Fig. 12. Intuition tells us that 2(a) has better resolution, but upon examining the spatial frequency domain, we need to give the question some more thought. The effective lateral spatial bandwidth in (d) is 74 percent greater than in (b). Thus - the lateral spatial resolution of the large curved wavefront (c) is potentially greater than the lateral resolution of the smaller focal region field (a)! To understand this better, notice that a lateral slice through the psfs looks very

much like a chirp function. The resolution must be recovered by processing similar to current coded excitation schemes . Similar to Jensen[69] we propose a matched filter technique, whereby a time-reversed copy of an RF psf (c) is used as a filter. When convolved with an RF image of a point target 12(c), a much narrower lateral distribution 13(b) is the result. The signal amplitude in 13(b) is also more than 20 times greater than that in 13(a), and 2.2 times greater than the focal gain. A marked improvement in visibility is seen when matched filtering is applied to simulated images of a cyst phantom, shown in Figs. 13(c) and (d). The matched filtering procedure is consistent with the strategy for the ideal or Hotelling observer, which is to whiten and match filter. In the nearfield, matched filtering without whitening performs well because the resulting image statistics are nearly white. It remains to be seen how robust the technique is to phase aberrations and other artifacts.

Regardless of whether spatial matched filtering is a practical technique, it shows an important point - that image processing can make information in the raw RF data more accessible to human observers, as seen in Fig. 13(d). Processing does not, however, increase the information content. The raw RF signal contains all the task-relevant data. Processing can at best maintain information content, and often, information will be lost. We should comment that beamforming schemes (at least for reception) can be viewed as either part of the system design or as part of the processing algorithms. Both views may have merit, and further research in this area could be pursued.

Before concluding, we comment on the general downward trend of detectability

with increasing axial distance as shown in Fig. 11(b). Diffractive losses mean that the farfield has less signal available than does the nearfield, and consequently the S/N_b decreases, dropping the overall detectability.

5.5. Discussion and Conclusions

Quanta (NEQ), a measure widely used by other imaging modalities. The detection signal to noise ratio is shown to be an integration of the generalized NEQ weighted by the spectral variance of the target signal (rather than the squared magnitude of the Fourier transform of the signal, as is the case with other modalities). This reflects that ultrasound systems are sensitive not to the magnitude of the medium parameters but rather depend on the variance (spatial fluctuations) of these quantities. Moreover, when object statistics are white over the sampling domain, the frequency-dependent weighting of $GNEQ$ becomes constant, and a target-independent picture of detectability can be summarized in a single number. The resulting framework behaves similarly to the quadratic task, is intuitive, and is amenable to measurement and prediction of system performance. The theory has been used to predict the information content of the ultrasonic beam at various field points. Similar analysis of location-dependent $IGNEQ$ may prove useful for evaluating the effectiveness of beamforming strategies. Currently lesion detectability is assessed with time-consuming Monte-Carlo simulations, or by using phantoms with real systems. Backscattered power inside and outside the lesion are measured in ratio to give a detectability metric. The figures of merit presented here are instead based on statistical models and have a dif-

ferent intention: they give an upper bound on the information content available with a particular system design. The theory does not attempt to predict human observer performance. However, new strategies mimicking the ideal observer are revealed for processing RF data that could improve detection of lesions for human observers. Experimental characterization of ultrasound systems is planned in the near future.

(a)

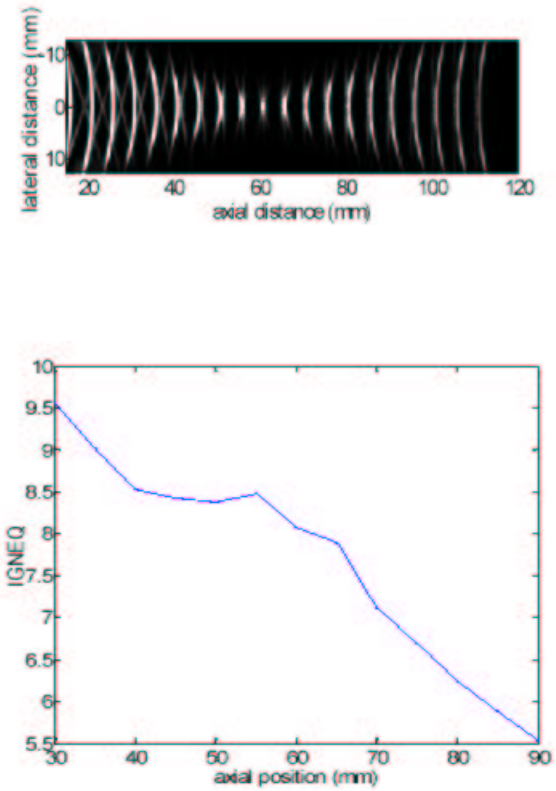

Figure 11. (a) Point spread functions due to 64 active elements of a 128-element linear array transducer of height 5 mm, element width l, and gap spacing of 0.1 mm, with fixed focus at 60 mm, and no elevation lens. An attenuationless medium was considered for simplicity. 2 cycles of a 3 MHz sinusoid weighted by a Hanning window were used to simulate the excitation pulse. The same function was used to simulate the electromechanical coupling impulse response of the transducer. (b) The normalized $IGNEQ$ values corresponding to field points along the beam axis. The larger the normalized $IGNEQ$ value the more informative the system is for the given field point.

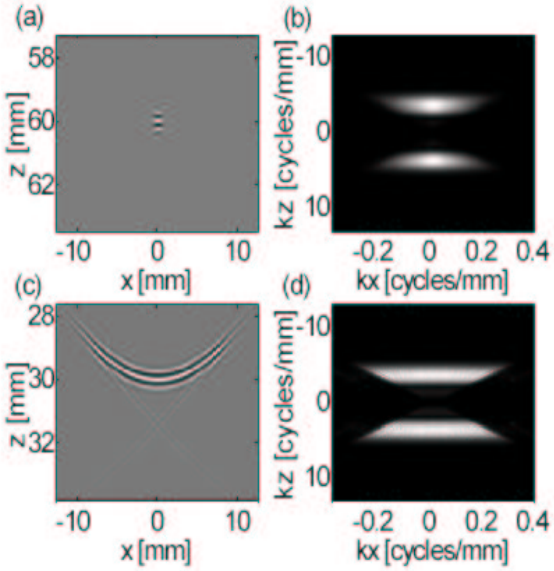

Figure 12. (a) Focal and (c) nearfield RF psfs due to a 3 MHz fixed focus linear array of height 0.5 cm, with 64 active elements of width l separated by distances of 0.1 mm. Azimuthal focus was 6 cm. The beam was electronically swept laterally across the point target located at (a) 6 cm and (b) 3 cm. (b) and (d) are the k-space (MTF) representations of (a) and (c) respectively.

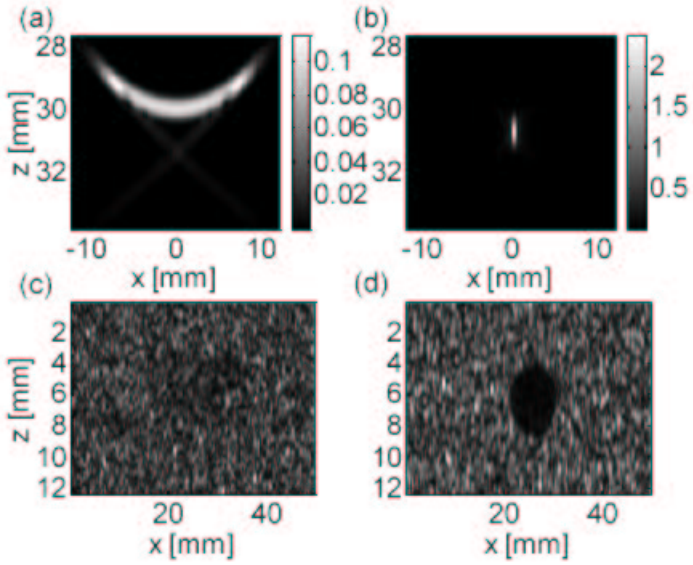

Figure 13. (a) The envelope-detected psf in 2(c). (b) The resulting image obtained after matched filtering 2(c) with a time-reversed replica filter. (c) Pre-filtered image of a lesion due to psf 2(c). (d) The corresponding post-filtered image.

6. IDEAL OBSERVER MODEL FOR ULTRASONIC DETECTION TASKS

6.1. Theory

6.1.1. Linear Systems Model and Assumptions

Medical ultrasound systems operate by transmitting pulsed beams of acoustic energy into the body and collecting the backscattered signal. Spatial variations in density and compressibility within the microstructure of the body act as scatterers. The RF or IQ echo signal $\mathbf{g}$ can be modeled as a linear system of the object function. IQ data is formed by mixing RF echo data with cosine and sine functions at the carrier frequency then low-pass filtering. B-mode signals discard phase information and merely detect the signal envelope, formed as $\mathbf{b} = \sqrt{\mathbf{g}_I^2 + \mathbf{g}_Q^2}$, where the squares and the square root act on the individual elements of the respective vectors. Alternatively, the B-mode data may be formed as either rectifying the RF data (i.e. $|\mathbf{g}|$) then low-pass filtering. Previous models of image quality for ultrasound systems have considered only the B-mode signal, which is a nonlinear function of the object and system. Consequently analysis was restricted to stringent assumptions. We are able to relax some of these assumptions by considering the pre-envelope (RF or IQ) signal. We are also interested in the pre-envelope signal to avoid losing information during B-mode processing. A solution to the small-signal acoustic wave equation in scattering media has lead to a linear systems model of ultrasonic imaging,[70] represented as $\mathbf{g} = \mathcal{H}f + \mathbf{n}$, where $\mathbf{g}$ is a vector of digitized echo signals, is a continuous to discrete integral operator, $f(\mathbf{x})$

is the object function, defined over spatial locations $\mathbf{x}$, and $\mathbf{n}$ is electronic noise. For a $N \times M$ digital image, $\mathbf{g}$ is a $NM \times 1$ RF column vector formed by concatenating the RF image columns. In this paper for simplicity we consider a 2-D system model, although a 3-D field model could be considered. Over small regions, often called isoplanatic patches, the system may be considered shift-invariant. We choose to analyze detectability over isoplanatic patches - an assumption that is not essential for the analysis, but greatly simplifies computation. Another reasonable approximation is that the object function can be considered discrete, so that the linear system can be written in matrix form

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n} \tag{90}$$

When correlations are much smaller than the image size, $\mathbf{H}$ may be approximated as a $NM \times NM$ block circulant convolution matrix.[23,71] This matrix notation will allow us accessibility to a wide range of statistical tools, matrix identities, and fast Fourier computational techniques. To model the elevation properties of the beam we would need to write the linear system as a linear superposition (in elevation) of planar convolutions, however, a 2-D model is presently justifiable - especially when considering targets with elevational depths greater than the elevational beamwidth. This simple approach helps us focus on optimization of the axial and lateral properties of the beam. Similar to Abbey,[35] we estimated system and noise parameters from a 7.5 MHz transducer using a Siemens Elegra Ultrasound scanner by imaging a graphite-agar phantom. The 2-D power spectrum was estimated from 10 independent images, and the pulse parameters were fit to a 2-D Gaussian function. The measured echo

SNR of these images was 47.0 dB (defined as the peak value of the signal power spectrum divided by the noise power level), and the estimated x- and y- pulse standard deviations were 0.08 and 0.15 mm respectively. The axial and lateral dimensions of a pixel are 0.054 mm by 0.12 mm respectively. Fig. 14 shows a representative schematic of the linear system model of image formation. Fig. 15 compares simulated and real data from a homogeneous phantom.

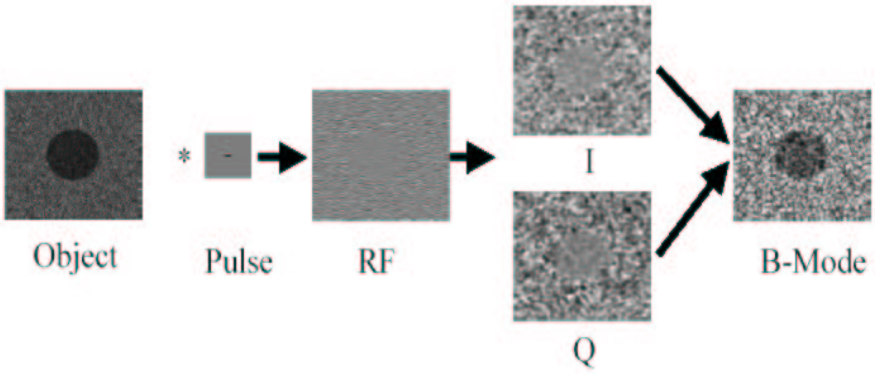

Figure 14. Model of the object function of a hypo-echoic lesion, the pulse, RF and IQ data and the B-mode image.

6.1.2. Signal Known Statistically

By modeling essential features of the system and the tissue, a statistical detection performance theory can be formulated. Although shape, echogenicity, heterogeneity, margin characteristics, and posterior acoustic shadowing are important diagnostic features in applications such as breast cancer sonography[72] we focus on echogenicity as shown in Fig. 14. Furthermore, although the role of breast sonography so far has been as an adjunct to mammography and as a means of discriminating lesions as

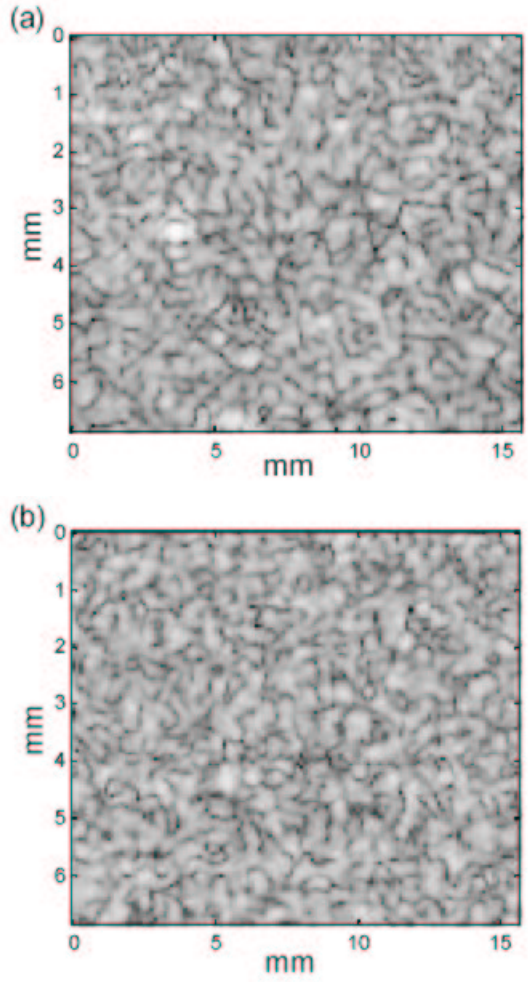

Figure 15. (a) B-mode data from a Siemens Elegra system scanning a homogeneous tissue-mimicking phantom (b) Simulated B-mode image.

benign or malignant, we choose to focus on ultrasound as a detection modality. This paper may furthermore be a stepping stone to evaluate performance for discrimination tasks (which are also binary classification tasks, like detection) since the theoretical descriptions may be very similar. We presently consider a signal known statistically (SKS) detection task. Aspects of the SKS task that are known are the noise power, lesion size, object contrast (echogenicity), shape, and location, as well as the system

spatial sensitivity function h. Unknown are the precise random realizations of the object function and noise processes. We assume that the object function can be modeled as a zero-mean Gaussian random process so that the echo signal is also a zero-mean Gaussian random process with distributions for signal present and signal absent hypotheses given as:

$$\begin{aligned} pr(\mathbf{g}|H_1) &= MVN(\mathbf{0}, \mathbf{K}_1) \\ pr(\mathbf{g}|H_0) &= MVN(\mathbf{0}, \mathbf{K}_0) \end{aligned} \tag{91}$$

In a manner similar to previous work[35,73,74,36] the signal present and signal absent covariance matrices have the form $\mathbf{K}_{data} = \mathbf{H}\mathbf{K}_{obj}\mathbf{H}^t + \sigma_n^2\mathbf{I}$ where $\mathbf{K}_{obj}$ represents the covariance of the object and represents the additive noise. We represent a lesion as an object region with a differing variance than the surrounding region (see Fig. 14). In general the signal present covariance matrix $\mathbf{K}_1 = \langle \mathbf{g}\mathbf{g}^t|H_1\rangle$ can be thought of as the sum of the background covariance and a differential signal covariance: $\mathbf{K}_1 = \mathbf{K}_0 + \Delta\mathbf{K}$ where $\Delta\mathbf{K} = \mathbf{K}_1 - \mathbf{K}_0$. For hypo-echoic lesions, $\Delta\mathbf{K}$ is negative. Throughout the paper we assume that the background process is wide-sense stationary.

6.1.3. Ideal Observer Test Statistic: The Log-Likelihood Ratio

The log-likelihood ratio is the decision function of the ideal observer. From 91, ignoring constants and data-independent terms, it may be written in terms of the RF data as[74,36,35]:

$$\lambda(\mathbf{g}) = \mathbf{g}^t(\mathbf{K}_1^{-1} - \mathbf{K}_0^{-1})\mathbf{g} \tag{92}$$

Note that the test statistic is quadratic in the data. When a SKE detection task was considered, the test statistic was linear in the data. One interpretation of the ideal observer strategy is to decorrelate the data with respective signal present and signal absent inverse covariance operations $\mathbf{K}_1^{-1/2}$ and $\mathbf{K}_0^{-1/2}$, then take integrate the intensity of the resulting images, and subtract. Note that the ideal observer test statistic cannot generally be obtained from the envelope signal, since the decorrelations cannot be performed after the phase information contained within $\mathbf{g}$ is discarded.

6.1.4. Monte Carlo Methods and Power Series

Abbey[35] introduced a 2AFC Monte Carlo method for assessing the performance of the ideal observer. A key contribution was a way of computing the inverse covariance matrices. The signal absent covariance is stationary, and may be approximated as a convolution operator - thus may be inverted quickly using Fourier techniques. The signal present covariance is however non-stationary, greatly complicating the inversion. Standard matrix inversion techniques are impractical given the size of the data covariance matrices considered (16,384 x 16,384 for the 128 x 128 images considered below). To accomplish the inversion, a power series was used. For a matrix $\mathbf{A}$,

$$(\mathbf{I}-\mathbf{A})^{-1}=\sum_{n=0}^{\infty}\mathbf{A}^n \tag{93}$$

The series is convergent when the eigenvalues of $\mathbf{A}$ are between -1 and 1. The inverse signal present covariance may be written as

$$\mathbf{K}_1^{-1}=\mathbf{K}_0^{-1/2}\left(\mathbf{I}+\sigma_{obj}^2\mathbf{K}_0^{-1/2}\Delta\mathbf{K}\mathbf{K}_0^{-1/2}\right)^{-1}\mathbf{K}_0^{-1/2} \tag{94}$$

The inverse operation may be implemented using 93 with $\mathbf{A} = -\sigma_{obj}^2 \mathbf{K}_0^{-1/2} \Delta \mathbf{K} \mathbf{K}_0^{-1/2}$. One never actually deals with these very large covariance matrices to implement the calculation of the test statistic. Instead $\mathbf{K}_0^{-1/2}$ and $\mathbf{H}$ can be implemented through Fourier techniques to be discussed shortly. The power series generally converges quickly for low and moderate contrasts.

The disadvantage of Monte Carlo methods is that one must generate a large number of simulated images which may be computationally burdensome. If one is doing this already for a parallel observer study, then this technique is a natural method. If this is not the case, quicker analytical predictors of performance may be used, as will be described below.

6.1.5. Pre-Envelope Deconvolution

The ideal observer processes raw data in an optimal way to make a decision - its performance is only limited by the diagnostic information content in the echo signal. Although a human observer model is needed to optimize display level processing, algorithms that mimic the ideal observer may prove useful. Truncating the power series expansion, a first-order approximation to the ideal observer test statistic (Eq. 2.9 of Ref.[35]) reveals that the ideal observer strategy is to deconvolve the data, with the convolution operation $\mathbf{W} = \mathbf{H}^t \mathbf{K}_0^{-1}$. Deconvolution followed by envelope detection transforms the raw data into a form that may be easily interpreted by a human observer, and has shown to improve human observer performance over standard B-mode detection.[35] Recent coded excitation systems have successfully used deconvolution

processing to gain significant advantages in the penetration of ultrasound systems.[62,18] Our detection level framework adds the potential to optimize penetration-resolution tradeoffs.

6.1.6. Statistical Distributions of the Log-Likelihood

We are interested in calculating the SNR of the ideal observer analytically, rather than by using Monte Carlo techniques. To do so, we need to know the moments of the log-likelihood ratio. These can be obtained from the characteristic function[75] of the test statistic, viewed as the Fourier transform of the probabilities $pr(\lambda|H_1)$ or $pr(\lambda|H_0)$. Consider the characteristic function for the hypothesis H_j (see Eqs. 17.22-17.24a of Ref.[65]):

$$\psi_j(\xi) = \int exp^{i\xi(\mathbf{g}^t\mathbf{J}\mathbf{g})} \frac{e^{-\frac{1}{2}\mathbf{g}^t\mathbf{K}_j\mathbf{g}}}{(2\pi)^{NM/2}\sqrt{det\mathbf{K}_j}} = \frac{1}{\sqrt{det(\mathbf{I} - 2i\xi\mathbf{K}_j\mathbf{J})}} \tag{95}$$

where $\mathbf{J} = \mathbf{K}_1^{-1} - \mathbf{K}_0^{-1}$. The characteristic function can be used to calculate the moments of the log-likelihood ratio. We may also use the identity in Eq. 17.13 of[65] to write the denominator of 95 as:

$$det(\mathbf{I} - 2i\xi\mathbf{K}_j\mathbf{J}) = exp\left[\sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}(2i\xi)^m tr\left[\left(-\Delta\mathbf{K}\mathbf{K}_0^{-1}\right)^m\right]\right] \tag{96}$$

The advantage of this form is that the series should converge quickly for small contrasts $\Delta\mathbf{K}$. The traces are computationally straightforward to compute by means of the 2DFFT, and represent the moments of the log-likelihood ratio. From a second order truncation of 96, the characteristic function is normal and the signal present distribution density is approximated by a normal distribution (Eqs. 17.26 and 17.27

of Ref.[65])

$$pr(\lambda|H_j) = MVN\left(tr[\mathbf{K}_j\mathbf{J}], 2tr[(\mathbf{K}_j\mathbf{J})^2]\right) \tag{97}$$

The normal approximation may alternatively be derived from the central limit theorem without the low contrast approximation for target areas that contain a large number of post-whitened speckle spots. From this point of view, the ideal observer test statistic is a sum of whitened pixel intensities - and each pixel may be regarded as a random process. The test statistic represents a sum of many independent random processes when the target region contains a substantial (e.g. ¿30) number of whitened speckle spots. The main point now is that the moments from 97 may be used now to calculate ideal observer detection performance.

6.1.7. Detection SNR for SKS Tasks

From the above analysis, the test SNR of the test statistic may be calculated from 97 as:

$$SNR^2_{Ideal} = \frac{\{tr[\mathbf{JK}_1] + tr[\mathbf{JK}_0]\}^2}{tr\left[(\mathbf{JK}_1)^2\right] + tr\left[(\mathbf{JK}_0)^2\right]} \tag{98}$$

When pixels are separated by a correlation distance, covariance matrices are approximately diagonal, and Eq. (98) reduces to Eq. (33) of the Ref..[29] Our approach allows for a more general correlation structure. The power series 93 and 94 can be used to evaluate the traces. A second-order (low contrast) approximation to the ideal observer signal-to-noise ratio is

$$SNR^2_{Ideal,LowContrast} \approx tr\left(\mathbf{K}_0^{-1}\Delta\mathbf{K}\mathbf{K}_0^{-1}\Delta\mathbf{K}\right) \tag{99}$$

For higher contrasts, more terms may be retained in the series expansion. Clarkson and Barrett introduced another SNR (Eq. 5.61 of Ref.[36]) related to the Bhattacharyya distance, a metric quantifying the "distance" between two general probability density functions. Reassuringly, a second order approximation to the Clarkson-Barrett SNR discussed in Ref.[36] converges to 99. In the next sections we focus on ways of analytically and numerically evaluating this important expression, and explaining how it can be used for system optimization.

6.1.8. WGN Object Model

If a white Gaussian Noise (WGN) object model is assumed for the signal and the background, the signal present object covariance may be written as $\mathbf{K}_{obj} = \sigma^2_{obj}(\mathbf{I} + \mathbf{S})$, where $\mathbf{S}$ is a diagonal matrix with elements rastorized from the signal variance template

$$s(\mathbf{x}) = \begin{cases} \Delta\sigma^2/\sigma^2_{obj} & \mathbf{x} \text{ inside lesion} \\ 0 & \mathbf{x} \text{ outside lesion} \end{cases} \tag{100}$$

The non-zero elements S_i of $\mathbf{S}$ thus represent a patch of differing variance in the object (see Fig. 14), and are negative for hypo-echoic lesions. They are related to the object contrast factor (OCF) for ultrasound defined by Insana and Hall[34] as the relative change of standard deviation: $OCF = \sqrt{S_i + 1} - 1$. The signal covariance matrix may be written as $\Delta\mathbf{K} = \mathbf{HSH}^t$ and the covariance of the signal present matrix can thus be written as:

$$\mathbf{K}_1 = \sigma^2_{obj}\mathbf{H}(\mathbf{I} + \mathbf{S})\mathbf{H}^t + \sigma^2_n\mathbf{I} \tag{101}$$

6.1.9. Fourier Techniques

The test statistic 92 and analytic expressions for the ideal observer SNR may be evaluated with Fourier techniques. We represent the 2-D Fourier transform in matrix form as $\mathbf{F}$, and write $\mathbf{K}_0^{-1} = \mathbf{F}^{-1}\mathbf{D}_{\mathbf{K}_0^{-1}}\mathbf{F}$ and $\Delta\mathbf{K} = \mathbf{F}^{-1}\mathbf{\Phi}_{\mathbf{\Delta K}}\mathbf{F}$. Because is the covariance matrix of a stationary random process it is block-Toeplitz and can be approximated as block circulant.[23] The discrete Karhunen-Loeve transformation is simply the 2-D discrete Fourier transform $\mathbf{F}$, hence the matrix $\mathbf{D}_{\mathbf{K}_0^{-1}}$ is diagonal with elements equal to the power spectrum of the background and noise processes $|H(\mathbf{u})|S_b(\mathbf{u})+S_n(\mathbf{u})$ rastorized into a vector where $S_b(\mathbf{u})$ is the power spectral density of the background object texture, and $S_n(\mathbf{u})$ is the noise power spectrum. We can also write $\Phi_{\Delta K} = \mathbf{D_H}\mathbf{\Phi_S}\mathbf{D_H^*}$, where is block-circulant and Hermitian, and is the diagonal matrix of the eigenvalues of H. The eigenvalues are rastorized elements of $H(\mathbf{u})$, the 2-D Fourier transform of the system psf $h(\mathbf{x})$.

As an example of the Fourier technique, consider calculating a scalar term $\mathbf{g}^t\mathbf{K}_0^{-1}\mathbf{HSH}^t\mathbf{K}_0^{-1}\mathbf{g}$ in the expansion of the test statistic (5) for a Monte-Carlo study. We would first write the term as $(\mathbf{Fg})^t\mathbf{D}_{\mathbf{K}_0^{-1}}\mathbf{D_H}\mathbf{FSF}\mathbf{D_H^*}\mathbf{D}_{\mathbf{K}_0^{-1}}\mathbf{Fg}$. To implement this, we would take the 2D RF image $g(x)$, take a 2D-FFT, multiply by the transfer function $H^*(\mathbf{u})/(|H(\mathbf{u})|^2S_b(\mathbf{u})+S_n)$, take an inverse Fourier Transform, multiply the variance template $s(x)$ element-by-element, take another 2D-FFT, multiply by the transfer function $H^*(\mathbf{u})/(|H(\mathbf{u})|^2S_b(\mathbf{u})+S_n)$, then multiply against the Fourier transform of the image $g(x)$ and sum the elements of the resulting matrix. A recursive algorithm

can be used to implement the full expansion of the test statistic. Fourier techniques will be used to obtain closed-form expressions for ideal observer detection performance.

6.1.10. Ideal Observer SNR for Low Contrast Lesions

Using Fourier techniques, and assuming a WGN object model, the ideal observer SNR in the low contrast limit, Eq. (13) then becomes

$$SNR^2_{Ideal,LowContrast} \approx tr\left[\mathbf{D}_{GNEQ}\mathbf{\Phi_S}\mathbf{D}_{GNEQ}\mathbf{\Phi_S}\right] \tag{102}$$

where $\mathbf{D}_{GNEQ}$. The elements of this diagonal matrix are the elements of the 2D $GNEQ$, or Generalized Noise Equivalent Quanta [18]

$$GNEQ(\mathbf{u}) = \frac{|H(\mathbf{u})|^2}{|H(\mathbf{u})|^2 S_b(\mathbf{u}) + S_n(\mathbf{u})} \tag{103}$$

rastorized into a vector. The Noise Equivalent Quanta used in photon modalities is given as $NEQ(\mathbf{u}) = G^2 MTF(\mathbf{u})/NPS(\mathbf{u})$, where G is the large area contrast transfer, MTF is the modulation transfer function and NPS is the noise power spectrum. NEQ used for photon imaging modalities has historical origins with Shaw[24] and others. NEQ represents the spatial frequency-specific density of quanta at the input of an ideal detection system that would yield the same output noise as the real system under evaluation. Generalized NEQ as described by Barrett and colleagues[68] provides provision for a stochastic background texture of the object. For ultrasound systems the $GNEQ$ quantity is a measure of the spatial frequency sensitivity of detecting a signal in a background texture and in the presence of electronic noise. In other words,

it is the fraction of speckle energy (rather than photon energy) that contributes to detection.

Because $\Phi_{\mathbf{S}}$ is Hermitian, the trace may be written as $\sum_{\alpha,\beta=1}^{NM}(D_{GNEQ})_{\alpha\alpha}|(\Phi_S)_{\alpha,\beta}|^2(D_{GNEQ})_{\beta\beta}$. Noting that Φ_S is a block-circulant convolution operator, we write this as a continuous integral over spatial frequencies:

$$SNR^2_{Ideal,LowContrast} \approx \int d\mathbf{u} GNEQ(\mathbf{u})\,[T(\mathbf{u}) * GNEQ(\mathbf{u})] \tag{104}$$

where * represents 2-D convolution over spatial frequencies $\mathbf{u}$. The integral is evaluated over the interval of sampling frequencies. It represents the ideal observer detectability in closed form in terms of the $GNEQ$ and the 'task'. The task is defined by the Fourier transform of the object signal variance template:

$$T(u) = |\mathcal{F}_{2D}\{s(\mathbf{x})\}|^2 \tag{105}$$

where, $s(\mathbf{x})$ is given by Eq. 101. For large lesions, the task $T(\mathbf{u})$ approaches a delta function with amplitude proportional to the lesion area A, and the convolution with $GNEQ$ in 104 results in only a slight blurring of the $GNEQ$ shape which we ignore. For lesions larger than several correlation lengths, the low contrast ideal observer SNR becomes

$$SNR^2_I = AC^2 \int \left[\frac{MTF^2(\mathbf{u})}{MTF^2(\mathbf{u}) + S_n(\mathbf{u})/|H|_o S_b(\mathbf{u})}\right]^2 d\mathbf{u} \tag{106}$$

where MTF is the modulation transfer function, defined as

$$MTF(\mathbf{u}) = |H(\mathbf{u})|/|H|_o \tag{107}$$

with $|H|_o$ defined as the maximum of $|H(u)|$. C is the lesion contrast given by $C = \Delta\sigma^2/\sigma^2_{obj} = (1 + OCF)^2 - 1$. The range of validity for the large area approximation will be tested in the numerical results section. The integral in **??** is evaluated over the spatial frequency domain out to the sampling frequency limits. Importantly, the integral in 106 is a target-independent figure of merit for characterizing detection performance of ultrasound systems. Eq. 106 has a form similar to the Smith-Wagner ideal observer SNR which can be written as,[29 34]

$$SNR^2_{Smith-Wagner} = \frac{A}{S_c} SNR^2_o C^2 \tag{108}$$

where S_c is the correlation area, $SNR_o = 1.91$ for Rayleigh statistics of the envelope signal,[28] and C is the small signal contrast. This convenient equation tells us that the greater the number of speckle-spots per lesion area, the better the detection performance. For the SKS ideal observer acting on the RF signal, the integral in Eq. 106 essentially replaces SNR^2_o/S_c in the Smith-Wagner theory. The new factor includes electronic noise, and the integrand represents the power spectral density of a pre-whitened (noiseless) echo signal. The integral in 106 is very much like an inverse correlation length of the pre-whitened data - or in other words the speckle spot density after pre-whitening - thus it reflects the pre-whitening potential.

An important difference between the SKS and SKE theories in previous literature[73] is that there is a quadratic dependence of the $GNEQ$ rather than a linear one. Thus, unlike the SKE theory for ultrasound,[73] the SKS theory predicts that the detection SNR^2 is proportional to the square of the contrast.

So far in this paper we have given 3 main approaches for computing the SKS ideal observer performance: (i) the Monte Carlo approach using the iterative power series as discussed in[35] (ii) The low contrast approximation, Eq. 104, and (iii) The large area, low contrast approximation, Eq. 106. All three have so far assumed fully developed speckle, local shift-invariance, and used the WGN object function model, however, more general object models could be considered. The Monte Carlo technique is the most accurate for large contrasts because the power series is not prematurely truncated, however, it requires extensive computation time compared to Eqs. 104 and 106. In the Numerical Results section, we wish to show that even for fairly small lesion sizes, and moderate lesion contrasts, Eq. 106 is a good approximation to the ideal observer. Eqs. 104 and 106 are the primary theoretical results of this paper.

6.2. Numerical Results

6.2.1. Accuracy of the Low Contrast and Large Area Approximations

We investigated the accuracy of the low contrast approximations discussed above. To do so, we compare the predicted detectability of Eqs. 104 and 106 with the performance of the Monte-Carlo approach using the power series 93 and 94. The Monte Carlo approach is the standard against which our approximations should be measured because it does not prematurely truncate the power series. For Monte Carlo simulations, 2000 pairs of signal present and signal absent IQ data were simulated using model parameters derived from measured ultrasound data in a manner similar to Abbey.[35] For each image pair in the 2AFC experiment, the ideal observer test

statistic was computed using the power series 96. The power series used up to 16 terms. The series was terminated when the contribution of a term fell below 0.001% of the test statistic. The image with the greater test statistic was scored as the lesion present image. The percent of correct choices was plotted over a range of contrasts and for three different lesion sizes. The 2AFC percent correct (PC) was then related to the Yes-No ideal observer SNR by[34]

$$SNR_{YN} = \sqrt{2}\Phi^{-1}(PC_{2AFC}) \tag{109}$$

The results are shown in Figs. 16 (hyper-echoic lesions) and 17 (hypo-echoic lesions). Note that for hypo-echoic lesions, the large area approximation, Eq. 106 over-approximates Eq. 104 advantageously for large lesions. Fig. 18 illustrates the representative lesion sizes. Error bars in the ideal observer SNR correspond to detectabilities evaluated at $PC + \delta PC$ and $PC - \delta PC$, where δPC is the standard binomial sample error , where N is the number of Monte Carlo trials.[34] As expected, for low contrasts, agreement is very good. Divergence is larger at higher contrasts, but for large lesions the SNR is so good that the probability of correct identification is essentially 100% - thus the disagreement may not be terribly important. The results show the predicted linearity with contrast. We investigated the linearity of the ideal observer SNR^2 with lesion area as predicted by Eq. 106. Illustrating this, Fig. 19 shows the remarkable linearity of SNR with lesion diameter even for lesion sizes equivalent to only a few correlation lengths (a 0.5 mm diameter lesion contains 3.3 lateral pulse standard deviations and 6.25 axial pulse standard deviations). Fig. 20

shows the predicted SNR of the ideal observer with decreasing electronic noise.

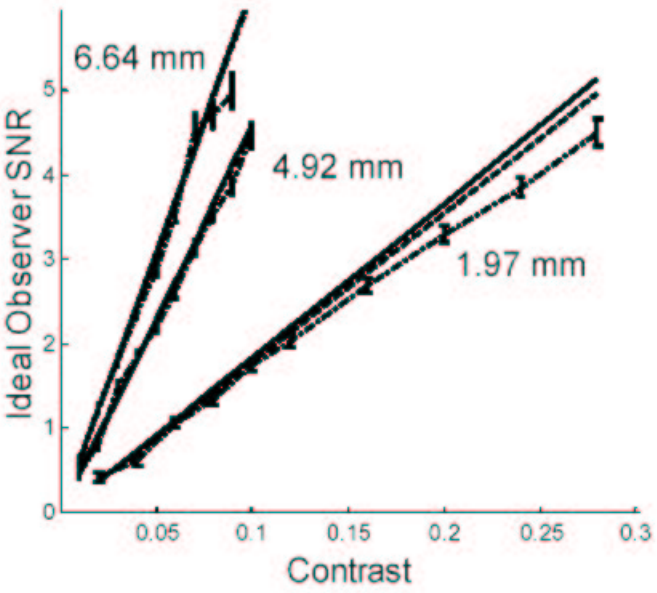

Figure 16. A test of the accuracy of the low-contrast approximation over a range of contrasts and for different lesion sizes for hyper-echoic lesions. Dotted line: Monte Carlo approach with iterative power series (gold standard), dashed line: Eq. 104, solid line: Eq. 106. Parameters were otherwise the same as outlined in section 6.1.1.

6.2.2. System Optimization

In high-noise environments, longer pulse lengths are sometimes desirable to improve signal energy and hence penetration. Longer pulses, however, degrade spatial resolution. We predicted ideal observer performance for a range of pulse lengths (with constant amplitude) for a 4.92 mm diameter lesion with contrast of -9.75%. In Figs. 21 and 22, we plot detectability curves for a range of echo SNRs. In Fig. 21, we used Eq. 106. Fig. 22 compares the Monte Carlo response and Eqs. 104 and 106.

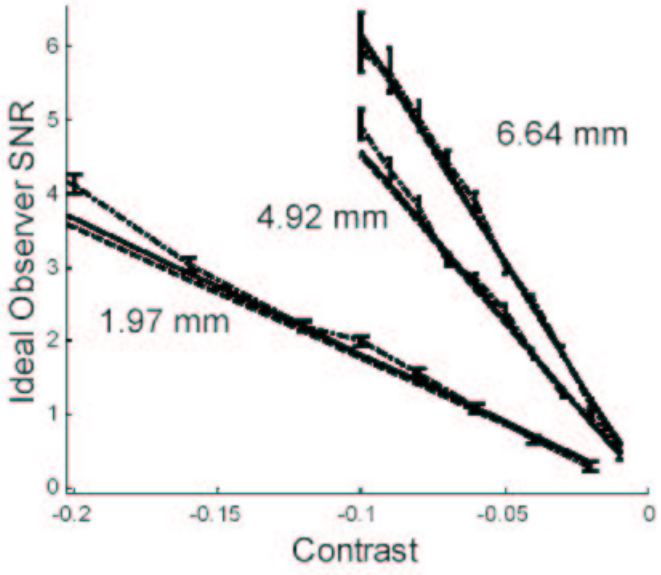

Figure 17. A test of the accuracy of the low-contrast approximation over a range of contrasts and for different lesion sizes for hypo-echoic lesions. Dotted line: Monte Carlo approach with iterative power series (gold standard), dashed line: Eq. 104, solid line: Eq. 106. Parameters were otherwise the same as outlined in section 6.1.1.

The approximations under-estimate the Monte Carlo results (due to truncation of the power series). Higher noise levels could correspond to deeper penetration depths, higher frequency excitations, or electronic shielding considerations. In high echo SNR situations, short pulses are seen to be superior, however, in high noise environments, longer pulses are seen advantageous. The optimum pulse lengths for varying echo signal to noise levels of 13, 9.5 and 7 dB are predicted to be 0.065, 0.1 and 0.13 mm respectively. For higher echo-SNR levels, shorter pulses are optimal, however, the length of the pulse will be wavelength limited.

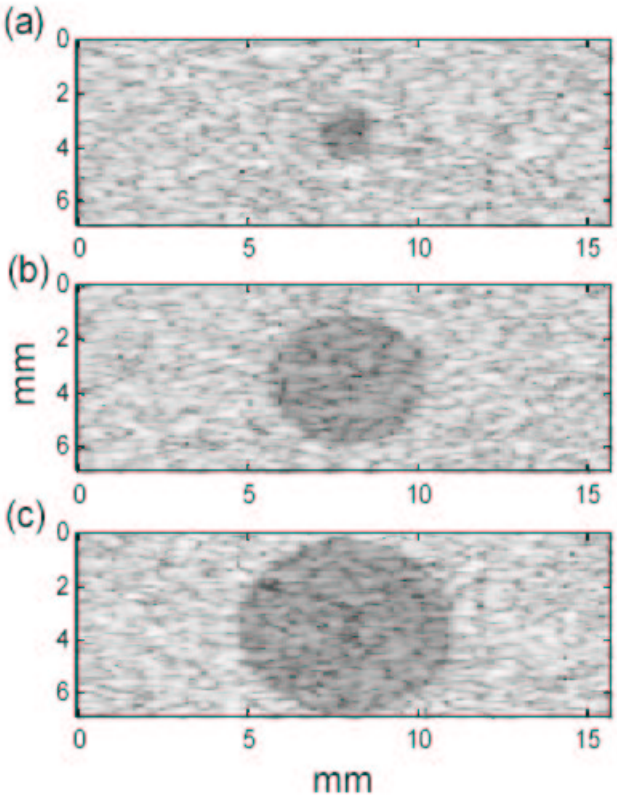

Figure 18. Representative lesions sizes used in Figs. 16 and 17: 1.97mm, 4.92 mm, and 6.64 mm diameters respectively. Contrast levels were purposely exagerated for visualization.

6.3. Discussion

The results of Figs. 16-20 suggests that the low contrast and large area SNR approximations of the ideal observer are reasonable to use - even for small lesions and moderate contrasts. For small lesions or large contrasts, the approximations are not necessarily within the error bars (confidence bounds) of the Monte Carlo results but are still remarkable close. For hypo-echoic lesions, the large area approximation 106

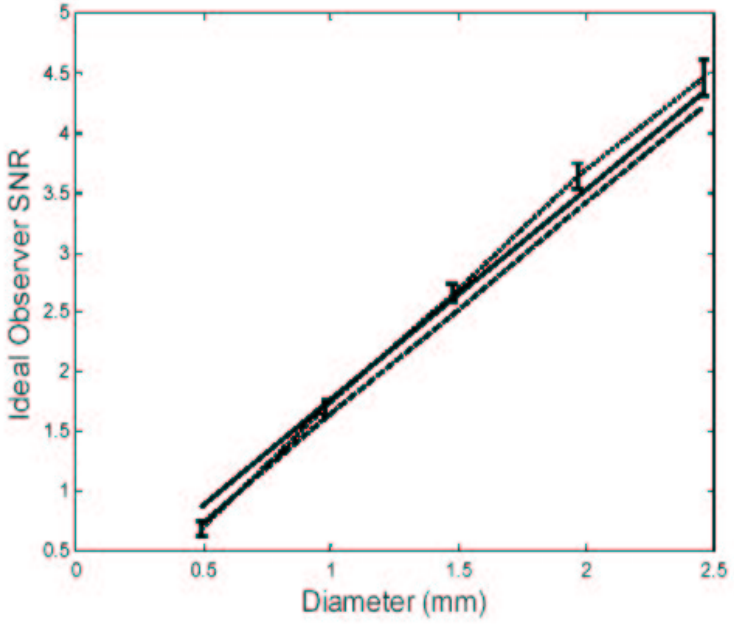

Figure 19. A test of the linearity of the ideal observer SNR with lesion diameter for a with -19% contrast (hypo-echoic) lesion. Dotted line: Monte Carlo (gold standard), dashed line: Eq. (18), Solid line: Eq. (20).

over-approximates Eq. 104 advantageously. The SNR expressions can be computed very quickly thus avoiding time consuming Monte Carlo methods. Eqs. 104 and 106 could be used as a quick first step in an optimization procedure. The Monte Carlo approach could then be used to fine-tune the optimization. The ability to quickly compute the ideal observer performance represents an important step in understanding how to optimize ultrasound systems for detection tasks.

Fig. 20 illustrates the impact of noise on the ideal observer. Electronic noise hinders the ability of the ideal observer to pre-whiten the data. In an ideal, noiseless system the ideal observer performance is only limited by finite sampling. Although

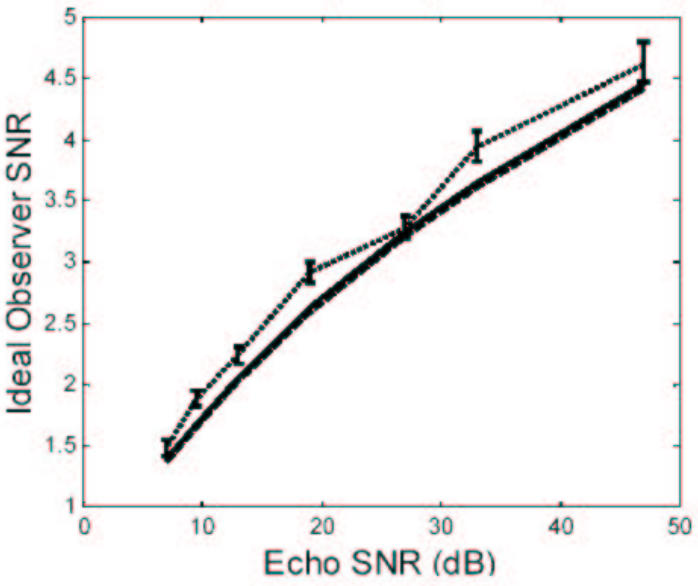

Figure 20. Ideal observer SNRs (dotted line: Monte Carlo, dashed line: Eq. 104 and solid line: Eq. 106 as a function of varying amounts of electronic noise for a 4.92 mm diameter lesion of contrast -9.75%. Parameters were otherwise the same as outline in section 6.1.1.

human observer performance (not investigated here) is likely to saturate and plateau with decreasing electronic noise (since human observers cannot pre-whiten well), the ideal observer's performance continues to improve. Any differences between the ideal observer and human observer suggest a role for image processing techniques such as pre-envelope deconvolution. This also means that anything one can do to reduce electronic noise or enhance signal energy may significantly enhance detection performance.

In the Smith-Wagner theory shorter pulses (larger bandwidths) are always better,

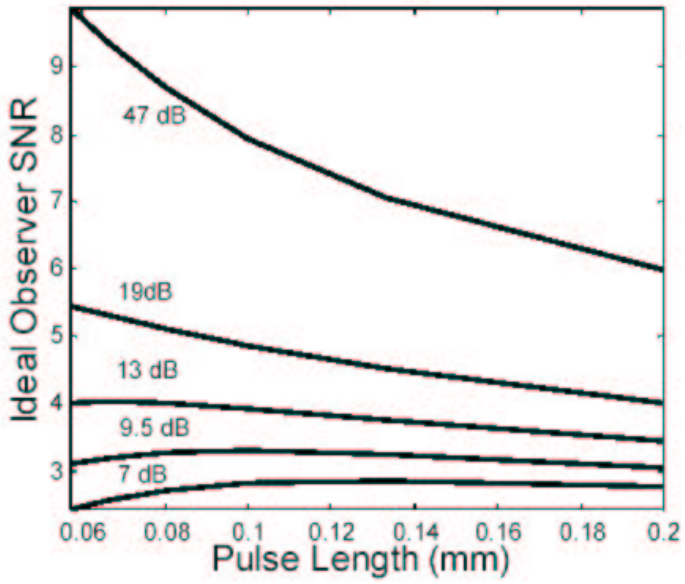

Figure 21. An example optimization study: ideal observer SNR as a function of pulse length for a range of echo signal-to-noise levels for a 4.92 mm lesion of contrast of -9.75%.

since in their derivations, electronic noise was regarded as being less important than speckle and thus ignored. Figs. 21 and 22 illustrate that the optimization point for pulse length is noise dependent. Fig. 22 shows that the Monte Carlo response and Eqs. 104 and 106 follow the same trend, even though the approximations underestimate the Monte Carlo results (due to truncation of the power series). The simple expression Eq. 106 can be used in a quick optimization procedure. If more accurate optimization is desired, the Monte Carlo technique with more terms of the power series can be used to fine-tune the maximum point.

In high noise conditions (e.g. high center frequency or deep tissues), there is a

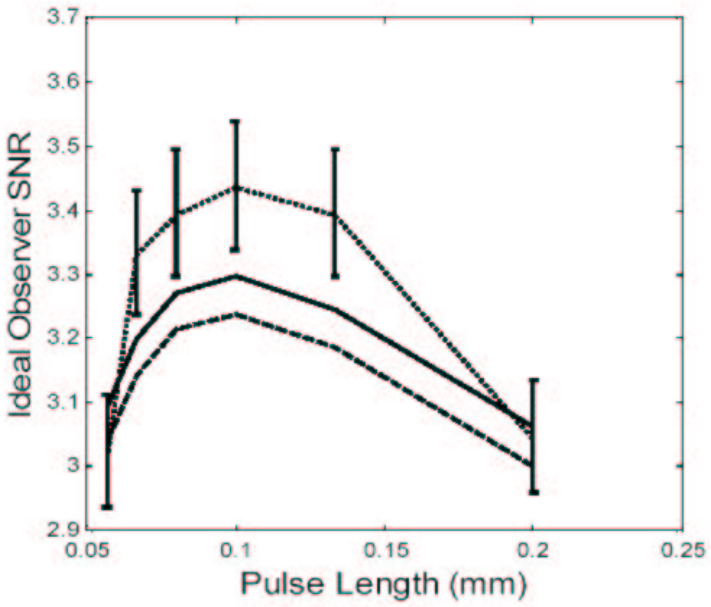

Figure 22. Pulse length optimization curve for 9.5 dB echo SNR, 4.92 mm lesion of contrast of -9.75%. Dotted line: Monte Carlo, dashed line: Eq. 104, solid line: Eq. 106.

tradeoff between penetration and spatial resolution. Longer pulses are more energy rich and thus penetrate more, but at the cost of resolution. Fig. 21 shows the optimal tradeoff for detection tasks. From Fig. 21, we see that the system performs moderately well even in high noise conditions. Depending on the situation, one may sometimes wish to sacrifice large ideal observer SNRs to optimize a parameter such as spatial resolution that is important for other tasks such as discrimination, and still perform very well on a wide range of detection tasks. The gains in detection performance from boosting SNR may depend on where a radiologist chooses to operate on the ROC curve. They will also be dependent on lesion size and contrast.

One way of improving spatial resolution could be to use higher transmit frequencies. Large time-bandwidth codes,[7618] could be used to boost echo SNR to compensate for increased attenuation - thus decreased penetration. This theory tells us how far we can push these limits without significantly sacrificing detection performance. It could also lead to the detection of much smaller lesions, thus reducing the number of false negatives.

6.3.1. Limitations and Extensions

It is important to discuss the limitations of our model. We have neglected aberration effects in our model, which may degrade detection and pre-whitening performance. A criterion for a good system design could be the relative robustness to aberration artifacts. Alternatively, de-aberration strategies may be useful. Our linear systems model is so far only 2-D. Although it is possible to extend the model to include the elevation properties of the beam, the analysis is not included here for sake of simplicity. 3D beam properties are a limiting factor when partial volume effects due to elevational clutter are important. 2-D analysis may be adequate when the elevation properties are of little importance.

The model has yet to include non-Gaussian statistics. Some investigators studying statistical models of ultrasound images often focus on the pixel statistics of B-mode images, and use Rayleigh, Rician, K-mean, Nakagami, and other distributions,[28].[23] Many of these can be generated from the envelope of a Gaussian RF model with spatially varying scattering amplitudes (variances) or means.

Our model has ignored background texture. Inhomogeneous background variability degrades detection and, for photon modalities, has been shown to influence system optimization.[32,77,31] Often more complicated object *pdf*s are unknown or difficult to obtain. Alternative observers could be studied that integrate measured power spectra, similar to techniques discussed by Burgess[77] for mammography. If background texture can be adequately modeled, Markov Chain Monte Carlo techniques[78] may be used to computationally evaluate the ideal observer performance. This approach can also model signal shape and location uncertainty.

The detection theory developed here may stand without the need for including background inhomogeneity. If there are regions of normal tissue texture that look like possible lesions, we would somehow desire to design the system hardware to suppress visibility of this texture while enhancing visibility of the true lesions - but for ultrasound systems how to do this is unclear. Human observers can sometimes distinguish between lesions and normal anatomy - based on features such as shape and contrast. The strategy then could be to reduce the number of missed true tumors by improving low contrast visibility - then rely on feature-based discrimination tasks to reduce the number of false positives. With this point of view, one design strategy could be to optimize the system low contrast detectability for the worst-expected (uniform) background echo SNR.

This paper says nothing about modeling human observer performance - a subject of future work. Models of human performance may greatly accelerate processing-level development.

An important application of this work is imaging breast cancer. Currently, ultrasound is typically performed after screening mammography - thus its role is not so much for detection but rather in discriminating between fluid filled cysts and solid masses, and discriminating between benign and malignant masses. Although this role could change to include detection should technology improve, the theoretical framework developed here may also provide a way for studying discrimination tasks.

6.3.2. Application to System Design

Instead of guessing how to trade-off system parameters optimally for detection tasks, we now have a rigorous and simple theory to guide the design. System design specifications could potentially be task-based. For example, one could specify a minimum lesion size and contrast for which some percent of successful detection is desired.

The theory in this paper could be used to explore several open avenues leading to improved system design. (1) O'Donnell[62] argued that ultrasound systems use only a fraction of the time-averaged energy that is possible from a regulatory point of view, and that coded excitation techniques could improve penetration while maintaining spatial resolution and while staying below peak transmit pressure limits. With the recent lift of these regulatory standards[79] the combination of codes and higher transmit amplitudes may offer significant gains. One may be able to use higher transmit frequencies or harmonic imaging,[80 81] to improve spatial resolution, while maintaining signal-to noise with codes. The theory here gives us a framework for how to push the limits of spatial resolution with the aid of codes. (2) Recent 1.75 or 2-D

array technology may greatly enhance the diagnostic performance of ultrasound systems.[22] (3) Novel transmit and receive beamforming schemes, and lateral codes[73] may also be an open area for additional improvement. (4) Spatial compounding with pre-envelope deconvolved B-mode images could be another way for reducing speckle variance while maintaining very fine spatial resolution. This will have the effect of reducing speckle, and averaging other artefacts. Realtime spatial compounding (without pre-whitening) has already shown promise in breast cancer applications.[82] A combination of coded pulses, higher transmit frequencies, pre-envelope deconvolution, and spatial compounding may offer significant image quality improvements and may be possible in realtime.

6.4. Conclusions

The equations and techniques in this paper offer a way of numerically evaluating the effectiveness of a system design. This paper makes connections to NEQ, used in photon imaging modalities, as well as other classic models of image quality in the literature. The generalized NEQ for ultrasound can be thought of as the fraction of speckle energy that contributes to the detection task. The task is seen to be the variance profile of the lesion, or its associated Fourier magnitude. Optimization of engineering tradeoffs such as resolution-penetration may be predicted using the theory. It is seen that detection tasks are robust to a significant amount of noise. It may thus be possible to use higher transmit frequencies to improve spatial resolution while maintaining contrast resolution. This may further be facilitated by codes

and higher transmit amplitudes. In this way smaller lesions may be identified, earlier, reducing the number of missed diagnoses. The relative merits of coded excitation techniques, and novel beamforming strategies may be evaluated and optimized for detection tasks. The framework should be able to accommodate nonlinear propagation, frequency dependent attenuation, and possibly background texture. Future work aims to experimentally test system optimization predictions. Finally, this analysis motivates RF deconvolution as a processing strategy that mimics the ideal observer, and hence may make information in the data more accessible to human observers.

7. EXPERIMENTAL VALIDATION OF STATISTICAL MODELS

Work in previous chapters has focused on deriving models for the ideal observer test statistic, and its associated performance metrics. Here we provide experimental evidence of the Gaussian nature of the statistics, validation of the linear systems model and its applicability, especially in regions away from the focal region (this is something which has not been done previously).

We also outline some techniques for estimating the ideal observer performance from experimental data, including appropriate error analysis.

7.1. Validation of Model Assumptions

7.1.1. Univariate Statistics

In our modeling we assumed that RF echo signals obey multivariate normal distributions with zero mean and covariance $\mathbf{K}$. Our first test of this assumption is to test the univariate statistics of RF echo signals from a homogeneous phantom, which should be Gaussian. We use a χ^2 test to assess how close the measured distribution is to a normal one.

We histogrammed RF pixel data from 128x128 patches in 32 independent scans of a homogeneous graphite-agar ultrasound phantom. Data was acquired using the Siemens Antares ultrasound scanner with a special Ultrasound Research Interface

(URI) that allows acquisition of the beamformed RF data before any nonlinear processing is performed. The echo signals were stored in a data file for transfer to a computer for offline analysis. The resulting histogram with 100 bins is shown in Fig. 23. A total of 507904 data points were included.

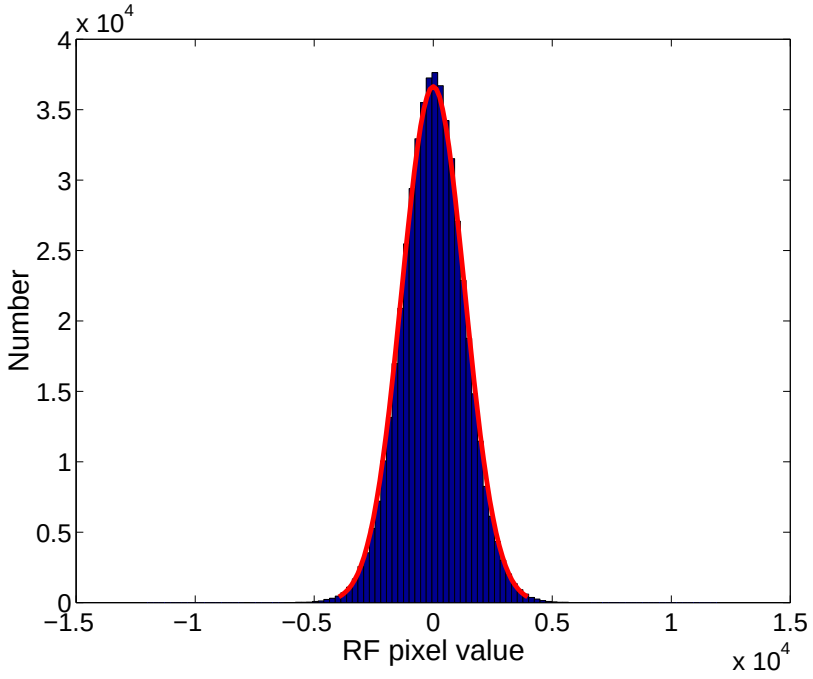

Figure 23. Histogram and fitted Gaussian curve of RF pixel values as taken from a homogeneous scattering phantom

Individual channel data are digitized 8-bit values ranging from 2^0 to 2^8. The beamformed RF data has a dynamic range greater than 8-bits because channels are summed together in a complex way. The RF data values are nonetheless integer values corresponding to digital values.

The Gaussian curve was estimated by estimating the mean and the variance from the data. The mean and variance were estimated as 1 and 1321 with 5% confidence

intervals of $[-2, 5]$ and $[1319, 1324]$ respectively.

To rigorously test the hypothesis that the observed data originates from an underlying Gaussian random process, we use the test statistic

$$\chi^2 = \sum_{k=1}^{K} \frac{(O_k - E_k)^2}{E_k} \tag{110}$$

where O_k and E_k are the observed and expected number of occurances in bin k respectively, and K is the total number of bins. This test statistic value has a χ^2 probability distribution.

For computed test statistic values smaller than some threshold, we decide the "null" hypothesis - i.e. the hypothesis that the observed data is normally distributed. For test statistic values greater than the threshold we reject the null hypothesis.

To reject the null hypothesis we choose the threshold value for which the false negative rate is less than 5%. The value of this threshold is given as the critical value $\chi^2_{0.05,\nu}$, where ν is the number of degrees of freedom. In our situation, since we used 100 bins, we had 100 observations. We used 3 constraints: the total number of samples, and the estimated mean and variance, hence we are left with $100 - 3 = 97$ degrees of freedom. The relevant critical value can be obtained by evaluating the inverse χ^2 cumulative distribution function with 97 degrees of freedom at a 0.95 probability level.

Using 100 bins over the interval [-3601,3569] the test statistic was computed to be 608.37. The critical value was computed to be equal to 120.99. Because the

computed test statistic is greater than the 5% threshold value we must conclude that it is improbable that the observed data is Gaussian. Although the shape looks quite Gaussian there are bins around zero that have slightly greater frequency than expected, as well as excess bins on either side of the Gaussian tail.

Does this mean that our ideal observer model assumptions (which assumed multivariate normal distributions) are incorrect? Not necessarily. It should be noted that a number of factors may have contributed to the deviation from the normal curve. The most likely factor is that the phantom was not completely homogeneous as we assumed. We needed the assumption of homogeneity to make use of ergodicity - thus data in different spatial locations in an image are used rather than requiring an ensemble of independent images. Inhomogeneity due to graphite clumping, air bubbles, or other factors can change the spatial variance profile of the phantom. As a result, when plotting a histogram of all such values, the distribution may not look normal - even though the underlying stochastic object and noise may be MVN with spatially varying variance profiles. Additional factors contributing to this may be tissue attenuation and shift variance. Echo signals attenuate with depth - thus introducing a spatially changing variance to the data. Additionally, shift-variance of the system point-spread function may have the same effect.

Unfortunately, we are left without any rigorous sense of whether the echo signals are MVN distributed or not.

7.1.2. Multivariate Statistics

Even assuming that we could construct perfectly homogeneous scattering phantoms, true validation of the multivariate normal assumption of RF data is difficult due to high dimensionality: the higher the dimensionality the more data that is required to get enough statistical data to form adequate histograms. Instead we study the joint distribution of 2-pixels at a time. In Fig. 24 we show the 2-D histogram of adjacent lateral pixels as an example.

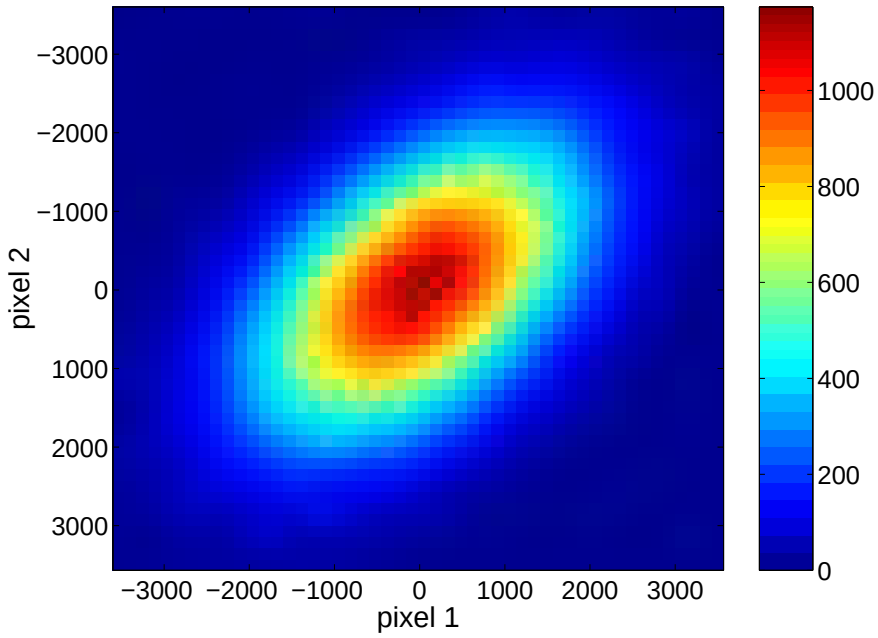

Figure 24. Observed 2-D Histogram of adjacent lateral RF pixel pairs as taken from a homogeneous scattering phantom. The colorscale represents the number of pixels that have RF value in the range of bin boundaries.

The RF data was the the same data used in the univariate analysis. The resulting 2-D histograms are elliptical in shape, centered about the origin, and the major

and minor axes are characteristic of the self- and cross-variance terms in the 2×2 covariance matrix. We test the MVN assumption- once again by a χ^2 test. To do so, we used a 50×50 binning scheme, and computed the χ^2 test statistic as above. The 2-D histogram of expected occurances is shown in Fig. 25, and was constructed as a multivariate normal distribution with mean and covariance estimated from the experimental data. The estimated mean and covariance of the data were $\widehat{\mathbf{p}} = [1, 2]^t$ and

$$\widehat{\mathbf{K}}_{\mathbf{p},\mathbf{p}} = \begin{bmatrix} 1.746 \times 10^6 & 0.845 \times 10^6 \\ 0.845 \times 10^6 & 1.745 \times 10^6 \end{bmatrix} \tag{111}$$

respectively.

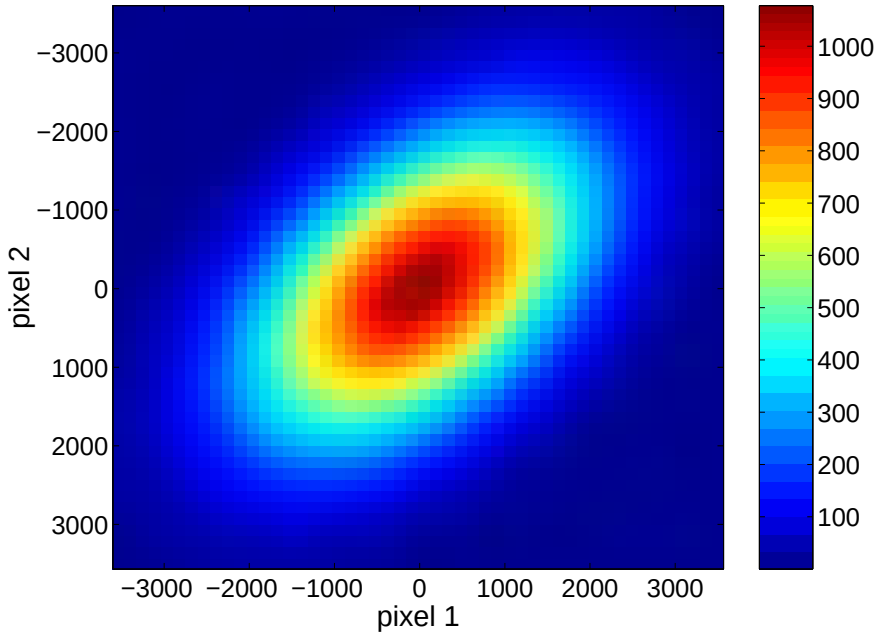

Figure 25. Expected 2-D Histogram of adjacent lateral RF pixel pairs as estimated from the mean and covariance of the experimental data.

To ensure an adequate number of samples per bin, we chose to perform the χ^2 test

in an elliptical window interior to the 5% isocontour line of the expected probability distribution. In this region, there were a total of 1393 bins. The χ^2 test statistic was computed as 2535. With 1 constraint on the total number of samples, 2 for the mean, and 3 for the covariance estimates (since $\widehat{\mathbf{K}}$ is Hermitian symmetric), we have $1393 - 6 = 1380$ degrees of freedom. Comparing the computed test statistic with the 5% critical value of $\chi^2_{0.05,1380} = 1475$, we see that we again must formally reject the null hypothesis that the test statistic is multivariate normal.

But, like the univariate test, we must account for phantom inhomogeneity, attenuation, and shift-variance as influences that may skew the normality of the observed data. These influences, however, do not rule out Gaussian object function and noise models. It is just that we are left without a formal way of validating the multivariate normal assumptions.

For fully developed speckle, where there are a large number of randomly positioned scatterers in a sample volume, by the central limit theorem, we maintain that a multivariate normal distribution for echo signals is a good model. Future efforts could be devoted to devising better experiments to validate these assumptions. There may be ways to minimize phantom inhomogeneity, attenuation, and shift-variance ... enough to demonstrate the MVN hypothesis in a rigorous way. For now the hypothesis will stand.

7.1.3. Covariance and Spatial Autocorrelation

We assumed that a 2-D ssf and object function were sufficient to generate the covariance matrix $\mathbf{K}$ characterizing the RF data. For a homogeneous scattering phantom, we hypothesized that the object function can be modeled by a 2-D white Gaussian noise process. We also assumed that over isoplanatic regions the system can be modeled as linear shift-invariant, and 2-D. In this case the covariance matrix $\mathbf{K}$ that characterizes the statistics of the RF data is stationary and described by the spatial autocorrelation function $R_{gg}(\Delta\mathbf{x})$. Here we test the assumption that the normalized autocorrelation function $\rho_{\mathbf{gg}}(\Delta\mathbf{x}) = R_{\mathbf{gg}}(\Delta\mathbf{x})/R_{\mathbf{gg}}(0)$ of the RF data can be adequately represented by $h(\Delta\mathbf{x}) * h(-\Delta\mathbf{x})$ normalized by its maximum value. Here we approximate h as the 2-D measured psf. Recall that it is actually more accurate to use the ssfs rather than psfs but measuring the ssf is difficult because it requires an ensemble of psfs and incremental depths. We hypothesize that using psfs in place of ssfs will be adequate over isoplanatic patches. Using the measured psf shown in Fig. 26 we thus compare $h(\Delta\mathbf{x}) * h(-\Delta\mathbf{x})$ shown in Fig. 27 with $\widehat{\rho}_{\mathbf{gg}}(\Delta\mathbf{x})$ shown in Fig. 28. To form the estimate $\widehat{\rho}_{\mathbf{gg}}(\Delta\mathbf{x})$, we used the 2-D unbiased estimator for the autocorrelation applied to RF data acquired from the Siemens Antares ultrasound system scanning a homogeneous graphite-agar phantom.

The RF autocorrelation is only roughly approximated using predictions from our 2-D linear systems model. To more accurately assess this, we need ssfs which are most easily obtained by having a good quantitative simulation model that can generate

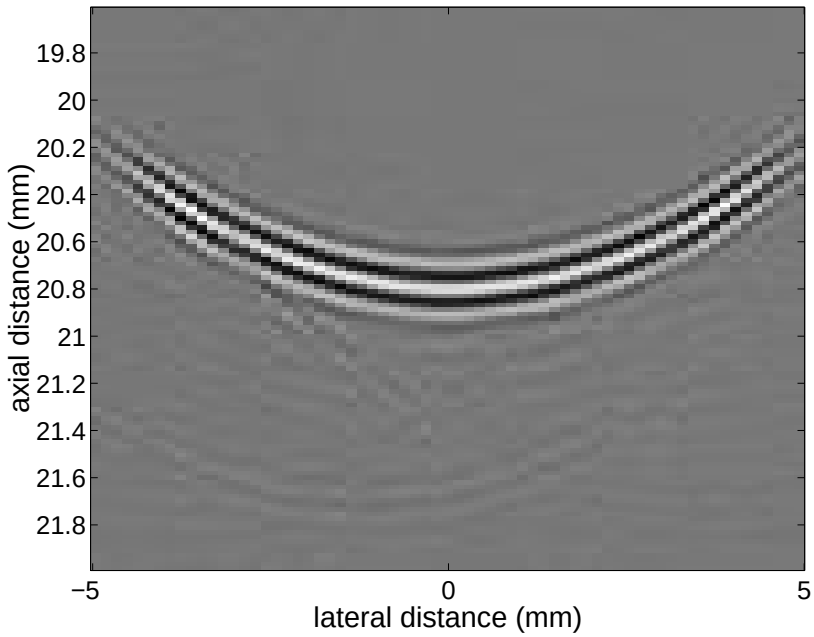

Figure 26. Measured nearfield psf

an ensemble of psfs at different depths. Currently our FIELD II simulations can only roughly approximate the observed system psfs. Additionally, shift-variance may contribute to loss of coherence - thus narrowing the estimated autocorrelation functions. We did not account for additive electronic noise, which we assumed was small - but this could nevertheless be another contributing factor to the narrower mainlobe observed in the estimate $\widehat{\rho}_{\mathbf{gg}}(\Delta\mathbf{x})$. Finally, the elevation dimension of the beam may contribute to some discrepancies between $\widehat{\rho}_{\mathbf{gg}}(\Delta\mathbf{x})$ and the normalized psf autocorrelation. All these contributions could be more carefully modeled in future work. For now, we prefer the simplicity of the 2-D model, and the local isoplanatic assumption.

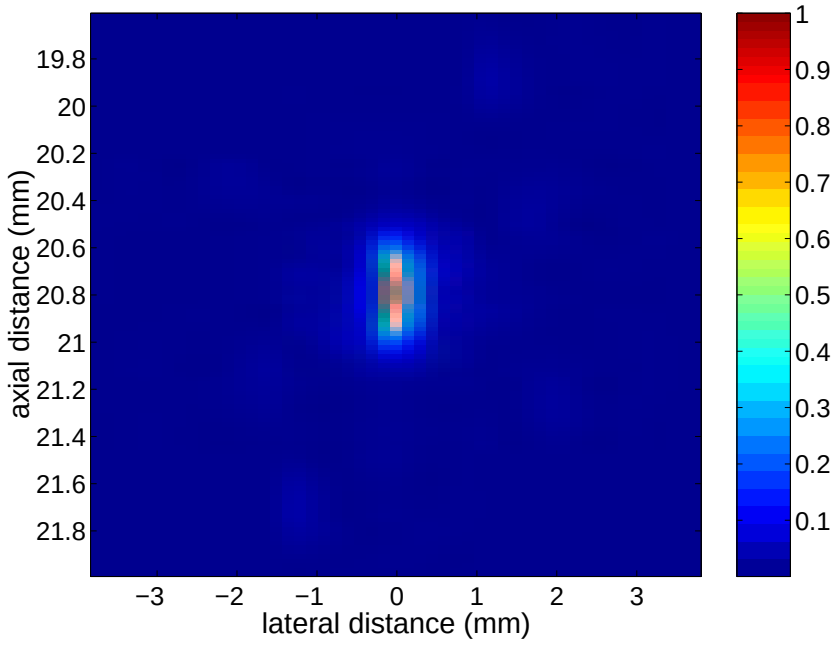

Figure 27. The autocorrelation magnitude of the measured nearfield RF point spread function.

7.2. Ideal Observer Test Statistic Normality

In the previous chapter, we assumed that the ideal observer test statistic could be assumed to be normally distributed by virtue of the cental limit theorem when many speckle spots were present in the lesion, or by virtue of a low contrast approximation when the contrast was sufficiently low. Here we test this approximation using simulated data. Echo data was simulated in the same way as described in the previous chapter, for lesions of radius 6 mm and 20% contrasts. 500 lesion-present / lesion-absent image pairs were generated. The test statistic for each lesion was computed and the test statistic values were histogrammed in 30 bins. χ^2 statistical hypothesis tests were performed on the observed data to test the hypothesis that λ was normally

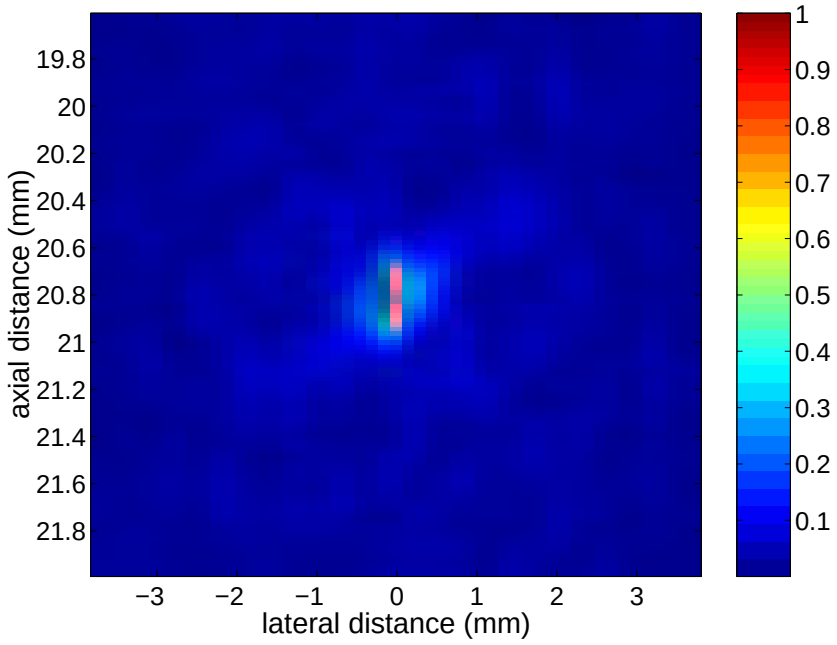

Figure 28. Estimated spatial autocorrelation magnitude from 7 independent patches of RF echo signals due to a homogeneous scattering phantom.

distributed. The resulting histograms for the signal present and signal absent cases along with their best Gaussian fits are shown in Figs. 29 and 30 respectively. For the signal absent case the χ^2 test statistic was computed as 32.6. Comparing this with the critical value 40.1, we need not reject the Gaussian hypothesis. Likewise for the signal present case, the computed χ^2 test statistic is computed as 32.4, which is less than the critical value 40.1, again confirming the Gaussian hypothesis.

The test statistic is more likely to be normally distributed when large lesions are considered so as to include many independent speckle spots. We consider a case where the lesion radius is only 1.85 mm and the contrast is again 20%. The test statistic for the signal present case no longer looks Gaussian and does not pass the χ^2 test (the

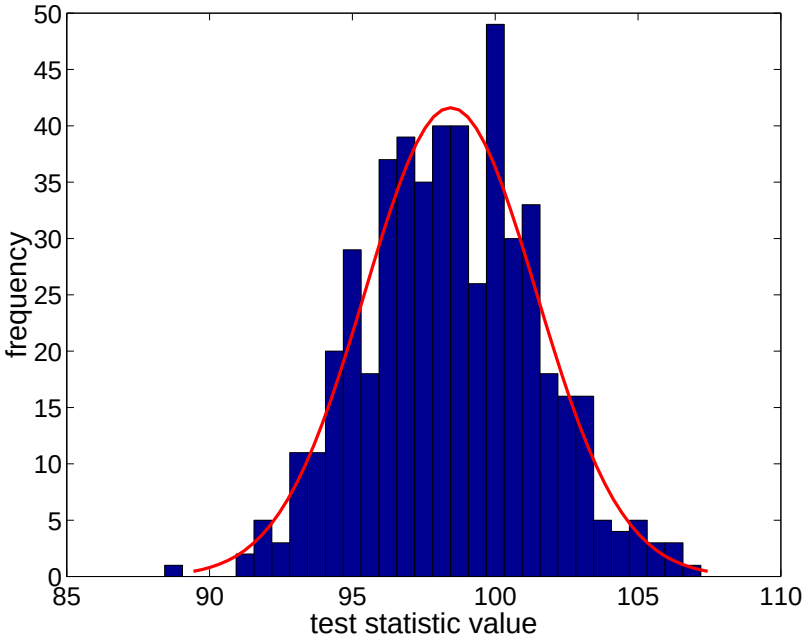

Figure 29. Histogram of signal absent test statistic values and best Gaussian fit.

calculated signal present χ^2 value of 70.1 is greater than the critical value of 40.1).

We should note that in situations where the test statistic is not normally distributed (for example, in small lesions), this does not mean that our analysis is invalidated. Rather it means that that SNR_I no longer rigorously parameterized the ROC curve. SNR_I may still be a valuable performance metric, and still maintains the Bhattacharya distance interpretation.

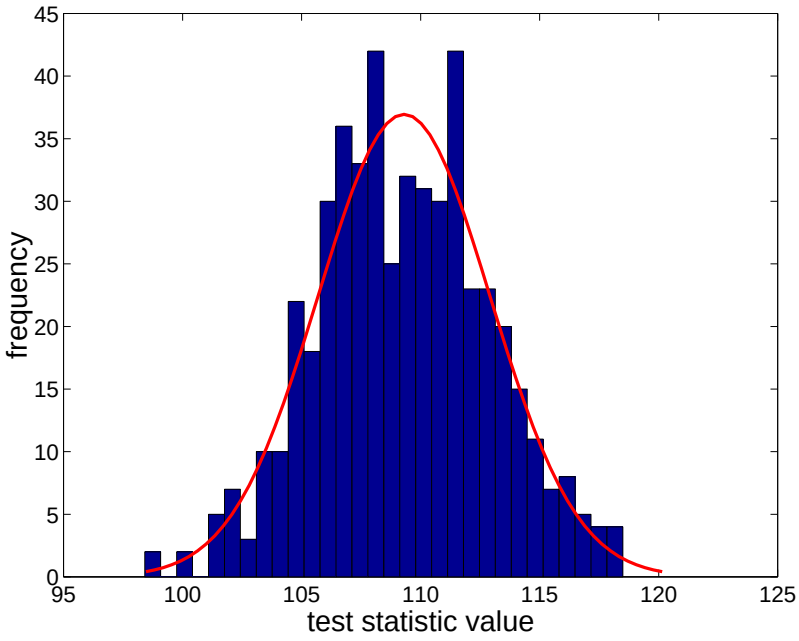

Figure 30. Histogram of signal absent test statistic values and best Gaussian fit.

8. UNFOCUSED IMAGING WITH SPATIAL CODES

8.1. Introduction

Geometric focusing, delay and sum array beamforming, and synthetic aperture reconstruction are examples of techniques used in medical ultrasound imaging systems to focus energy and attain high spatial resolution in sonograms. Synthetic aperture techniques accomplish focusing with multiple transmit-receive sequences with small apertures to simulate over time acquisition of a large spatial aperture that cannot be generated instantaneously as may be the case for array systems with limited channels, or for single element systems.

In this paper we discuss an alternative imaging technique that can be viewed as a generalization of synthetic aperture imaging. We propose a simple form of spatial coding of transmitted beams, and processing algorithms to accomplish decoding to recover spatial resolution. Similar to coded excitation techniques, where a digital code or chirp is transmitted in the axial direction,[62,18,83,76] we propose lateral or spatial encoding. Other groups have proposed similar ideas previously. Fink **??** proposed the idea of filtering echo signals with a time reversed echo from a bright scatterer. Additionally, Jensen and Gori[69] proposed that focusing can be accomplished by spatial matched filtering. Jensen[64] and Anderson[63] independently proposed the concept of adding lateral modulation to beams for the purpose of lateral flow estimation. We extend this idea to lateral coding with a baseband chirp. Our analysis adds theoretical background and explanation to previous work, and makes quantitative the potential

benefits. We call our technique 'unfocused imaging' or 'spatial coding' because there is a natural spatial encoding that occurs in the near and farfield (unfocused regions) of focused transducers - an observation that will be discussed shortly.

8.2. Theory: Ideal Observer

8.2.1. Task Performance

The technique of spatial coding is one example of a novel imaging technique that has stemmed from a much broader fundamental analysis of image quality assessment. The analysis gives specific predictions about the diagnostic information in the near and farfield of ultrasound beams. For this reason we discuss the salient features of the theory of task performance here.

Performance of medical ultrasound imaging systems is typically summarized by parameters such as spatial resolution, echo SNR and contrast resolution. There are tradeoffs among these parameters which are inherently application dependent. A more rigorous approach is to measure image quality by the quantitative performance of specific diagnostic tasks. In this paper we consider two types of tasks:

(1) detection of low contrast lesions and

(2) discrimination between anechoic cysts and high contrast hypo-echoic solid masses.

More emphasis will be given to detection tasks since more has been done on the topic in previous work.

In an information-theoretic paradigm, a system should be designed to maximize information content for the task at hand. Processing and display algorithms should then be devised to optimally convey the diagnostic information to the decision maker. Processing algorithms cannot add information to the RF echo signals - processing will always be either lossless or lossy - thus the maximum amount of information resides in the unprocessed RF echo data. We concentrate on optimizing information content of the raw signal - based on maximizing the performance of a Bayesian ideal observer (in the case of binary classification tasks). The ideal observer is characterized by the optimal scalar decision statistic - the log-likelihood ratio.[23]

$$\lambda = ln\left(\frac{pdf(\mathbf{g}|H_1)}{pdf(\mathbf{g}|H_0)}\right) \tag{112}$$

where $pdf(\mathbf{g}|H_1)$ and $pdf(\mathbf{g}|H_0)$ are the likelihoods of the RF echo data $\mathbf{g}$ conditioned on the signal present (H_1) and signal absent (H_0) hypotheses. It is optimal in the sense that the decision thresholding process gives the highest true-positive probability for a fixed false-positive level. Thus - if there is any processing to be done to the RF signal the ideal observer knows how best to do it. Our design philosophy is to optimize the performance of the ideal observer for diagnostic tasks and then attempt to design processing algorithms that attempt to mimic the ideal observer processing strategies.

8.2.2. Stochastic Data Models and Linear Systems

Stochastic models of echo signals are needed since the ideal observer has full knowledge of the data likelihoods. A linear systems approach to ultrasonic image formation has

been taken, such that the RF echo signal $g(t)$ can be written as a vector

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n}, \tag{113}$$

where **H** is the system matrix that maps objects into data space, **f** is the object function, and **n** is additive noise (all approximated as discrete in this case). **g** and **n** are $NM \times 1$ column vectors for a $N \times M$ pixel image. **f** and **n** may be taken as zero mean Gaussian stochastic processes. Consequently **g** is multivariate normal.

In much of this paper we assume that the image formation model can be adequately represented by a 2-D convolution model (in this case **H** is block circulant) as represented in Fig. (31). A model incorporating the 3rd dimension of the beam and object is possible but not necessary for our purposes. A lesion is represented as an area of differing object variance.

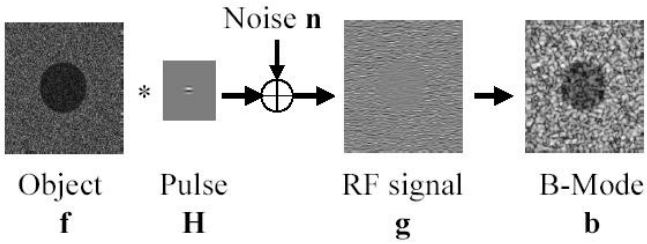

Figure 31. Linear system model of image formation

8.2.3. Ideal Observer Model for Ultrasonic Detection Tasks

Based on this linear systems model, we have developed an ideal observer model for ultrasonic low contrast detection tasks. The model is an extension of the theory

developed by Smith and Wagner in the early 80's.[29,28] The extension allows the inclusion of phase information by assessing the information content in the raw RF echo signals rather than the B-mode signal or focal I-Q signals. The Smith-Wagner approach in essence assumed that the speckle size was related to the size of the point-spread function - an assumption that is only true in the focal region. Our approach extends previous work to include regions away from the focal region.

The ideal observer theory may expand our understanding of what makes a good ultrasound system. It is based on rigorous mathematical groundwork, that we now probe using simulations and experiments in this paper. It has been shown using a power series technique and with a low-contrast approximation that the ideal observer decision function, the log-likelihood test statistic, takes the form[35,84]

$$\lambda = (\mathbf{K}_0^{-1}\mathbf{H}^t\mathbf{g})^t\mathbf{S}(\mathbf{K}_0^{-1}\mathbf{H}^t\mathbf{g}) \tag{114}$$

where $\mathbf{K}_0 = \langle \mathbf{g}\mathbf{g}^t | H_0 \rangle$ is the covariance of the signal absent echo data, and $\mathbf{S}$ is a diagonal matrix characterizing the lesion whose elements are zero when matrix locii correspond to points outside the lesion, and take values proportional to object contrast inside. The test statistic is quadratic in the data. The ideal observer decision strategy is thus to prewhiten the RF data by Wiener deconvolution, $\mathbf{K}_0^{-1}\mathbf{H}^t\mathbf{g}$, (more on this shortly), then integrate the intensity of the resulting signal in the region of interest.

The performance of the ideal observer for detection and discrimination tasks can be expressed in terms of SNR_I, the ideal observer signal-to-noise ratio. This quantity is not the same as the echo SNR. Instead it reflects the relative separation of

lesion present and lesion absent test statistic distributions cumulated from many independent echo signal acquisitions. For the task of detecting low-contrast lesions, a convenient expression for the ideal observer SNR has been derived as[84]:

$$SNR_I^2 = AC^2 \int d\mathbf{u} \left(\frac{MTF(\mathbf{u})^2}{MTF(\mathbf{u})^2 + 1/SNR_e}\right)^2 \tag{115}$$

where A is the lesion area, C is the lesion contrast, $MTF(\mathbf{u})$ is the Modulation Transfer Function,[23] $\mathbf{u}$ is a spatial frequency vector and SNR_e is the echo SNR. This expression links the features of the object and the system (including noise) to indicate how well the best possible observer performs. Performance is better when SNR_I is larger. This occurs when noise is low, system bandwidth is high. Large areas of high contrast are also easier to detect.

8.2.4. Depth Varying Detection Performance of a Focused Transducer

We thought it would be interesting to study the ideal observer performance over isoplanatic patches of a fixed f-number ultrasonic beam. To do so, we simulated psf's over increasing depths using FIELD II,[53] then calculated the SNR_Is. The probe we simulated was a VF10-5 linear array transducer from a Siemens Antares system. We specified the transmit and receive focus at 4cm. In what follows we use this probe unless specified otherwise. Nearfield images are taken to be around 2cm in depth. The psfs and corresponding ideal observer performance is shown in Fig. 31.

Because the psf is the smallest at the focus we expected the maximum performance to be there. Unexpectedly, it is not - the nearfield is actually best. We must conclude that there is substantial information in the nearfield that may be extracted

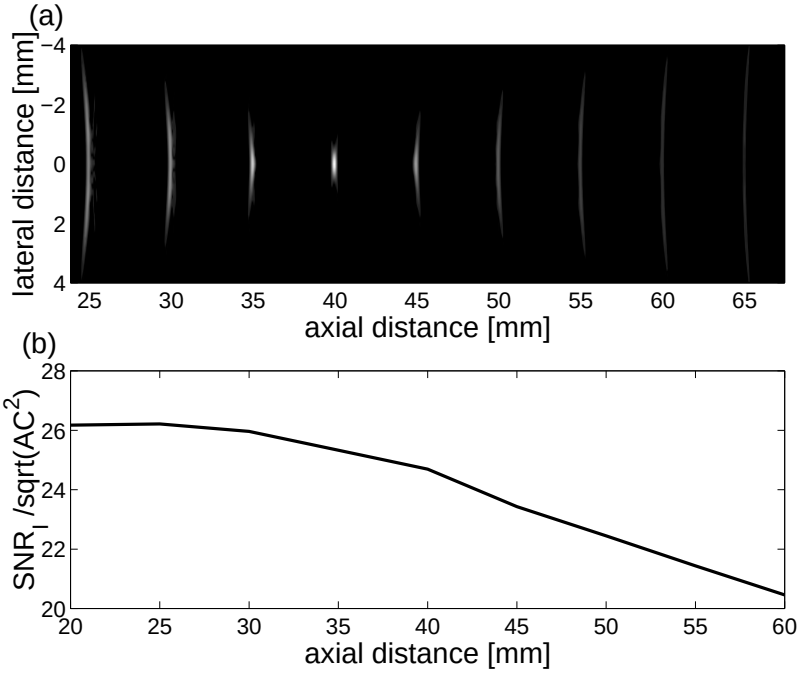

Figure 32. (a) B-mode psfs of scatterers (b) Corresponding SNR_I normalized by $\sqrt{AC^2}$ as a function of axial depth. Focal depth = 40 mm, F-number = 2.1, echo SNR = 6 dB.

by RF processing. In the following sections we explain the result and how to process the RF data.

8.2.5. Explanation of Unfocused Performance: Wavefront Curvature

To better understand the concept of unfocused imaging, consider both a nearfield and a focal region psf as shown in Fig. 33. Although the nearfield psf is more spatially distributed, the K-space 2-D Power Spectral Density $|H(\mathbf{u})|^2$ exhibits substantial lateral bandwidth - comparable to or even more than the focal region. This can be understood by considering that a horizontal slice through the nearfield psf is a chirp-like function with DC near the center. Thus the near (or farfield) delay and

sum beamforming or geometric focusing has provided a natural lateral chirp-encoding through *wavefront curvature*, similar to coded excitation techniques that are applied in the axial direction. It is BW which truly defines the spatial resolution potentially available for imaging.

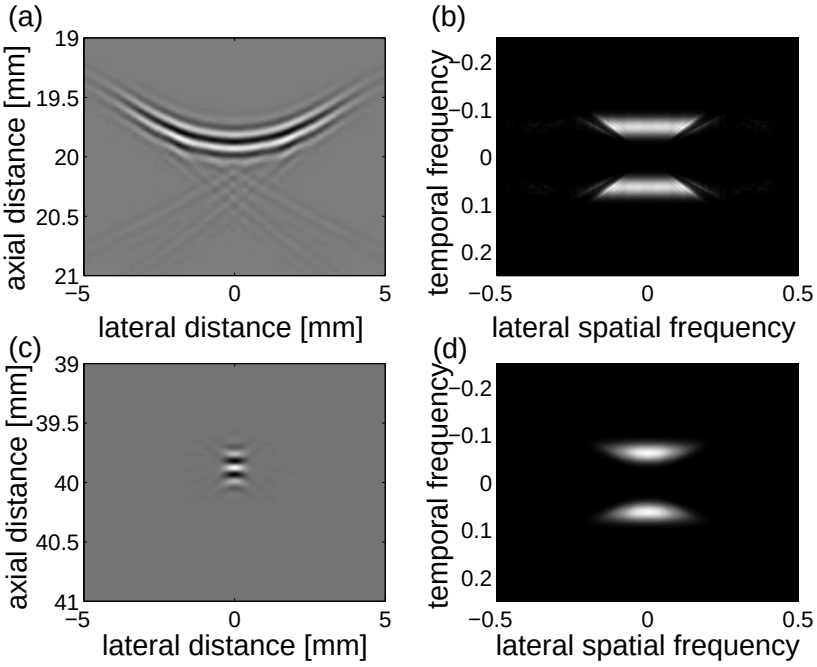

Figure 33. (a) Nearfield psf (b) K-space of nearfield psf (c) focal region psf (d) K-space of focal region psf. Parameters to be added.

8.2.6. Processing Algorithms to Recover Spatial Resolution

The ideal observer test statistic suggests that one should first filter the RF echo signals g with a filter $\mathbf{K}_{-}^{-1}\mathbf{H}^{t}$ then form an intensity image of the resulting signal. The test statistic is then simply computed by integrating this pre-whitened image over the lesion area. The filter $\mathbf{K}_{-}^{-1}\mathbf{H}^{t}$ can be viewed as a spatial Wiener filter. To see this, consider that both $\mathbf{K}_{-}^{-1}$ and H^{t} are block-circulant convolution operations for systems

that are linear shift-invariant in a local isoplanatic region, and thus are diagonalizable by a discrete Fourier transform. We note that that the eigenvalues of $\mathbf{K}_-$ are simply the power spectrum $|H(\mathbf{u})|S_f(\mathbf{u})+S_n(\mathbf{u})$ of the lesion-absent echo signal, where $S_f(u)$ is the power spectrum of the object function, and S_n is the noise-power spectrum. The continuous-time Fourier representation of $\mathbf{K}_-^{-1}\mathbf{H}^t$ is thus proportional to

$$W(\mathbf{u}) = \frac{H^*(\mathbf{u})}{|H(u)|^2 + \alpha\frac{1}{eSNR(\mathbf{u})}} \tag{116}$$

. with $\alpha = 1$ which a spatial Wiener filter.

A spatial matched filter is a special case of a Wiener filter in a noisy environment or when the regularizing parameter α is chosen to be very large. A time domain representation of the spatial matched filtering operation is

$$w(\mathbf{t}) = h(-\mathbf{t}) * g(\mathbf{t}) \tag{117}$$

where $g(\mathbf{t})$ is the RF echo signal, $h(-\mathbf{t})$ is the time-reversed point-spread function, and $w(\mathbf{t})$ is the filtered echo signal. Since the echo signals are truly data signals acquired over time, we consider g, h, and w as a function of vector-valued time, where the vector t has spatial correspondence defined by the imaging pulse sequence - more on this in section III.A.

8.3. Theory: Acoustics

At this point we turn to physics to quantitatively explain diagnostic information content in unfocused beams. We study two types of acoustic sources: focused transducers and point sources - both give analytic insights into the image formation procedure

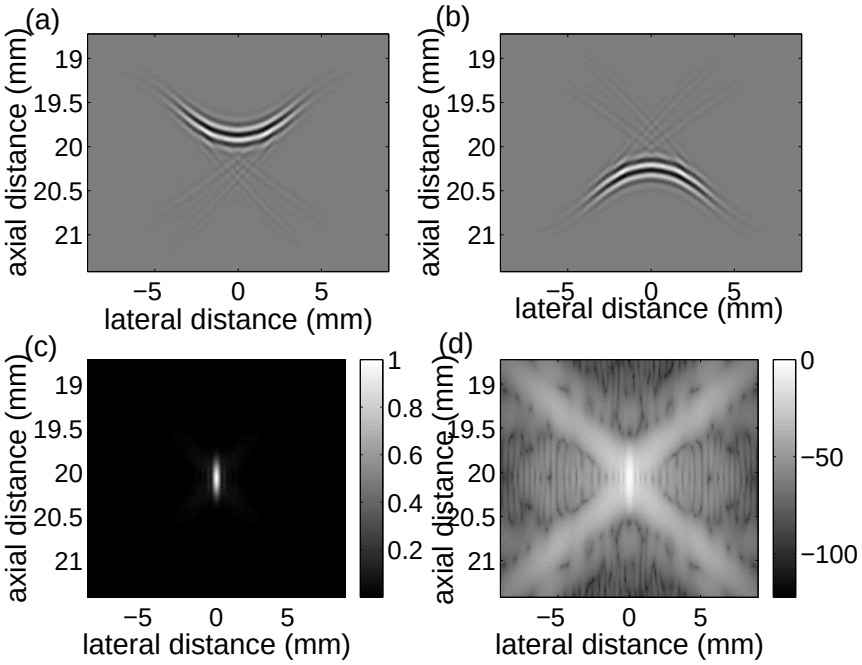

Figure 34. (a) Nearfield psf (b) matched filter (c) filtered image (envelope detected) (d) Log-scale spatial autocorrelation. Parameters to be added.

and the limits of spatial resolution. We assume that the sources are scanned laterally, and not in a sector format - although the theory could be expanded to analyze this.

8.3.1. Unfocused BW in Focused Transducers

We first concentrate on writing the equations of curved wavefronts from focused transducers. Our goal is to derive the lateral bandwidth of a transducer at axial depths at and away from the focus.

Linear Systems Model: To analyze the lateral bandwidth of a beam, we need a rigorous theoretical framework to accurately represent the diffractive propagation of ultrasound from a transducer. Here we cast the linear systems model of image

formation, Eq. 31 in continuous form, and add the physics necessary to understand spatial coding. The equation of ultrasonic image formation has been shown to be

$$g(\mathbf{t}) = \int h(\mathbf{x}, \mathbf{t}) f(\mathbf{x}) d\mathbf{x} + n(\mathbf{t}) \tag{118}$$

where g is the RF echo data, h is the spatio-temporal impulse response function, and f is the object function summarized by the spatially varying variance in the scattering function. The echo signal $g(\mathbf{t})$ is strictly speaking a function of time (time here can be considered a vector with structure characteristic of the pulse sequence used) rather than space, but there is typically a time-to-space mapping $\mathbf{x} = \mathbf{Bt}$ for us to associate the echo signal in a spatial context. (see**?** for details). The spatio-temporal impulse response can be written as a series of temporal convolutions,

$$h(\mathbf{x}, t) = h_y(t) * h_y(t) * v(t) * \frac{\partial h_a(\mathbf{x}, t)}{\partial t} * h_a(\mathbf{x}, t) \tag{119}$$

where h_y is the electro-mechanical impulse response of the transducer, $v(t)$ is the voltage excitation, and h_a is the acoustic impulse response of the transducer given by the Rayleigh integral

$$h_a(\mathbf{x}, t) = \frac{1}{2\pi} \int_S dS \xi(r) \frac{\delta(t - |\mathbf{r} - \mathbf{x}|/c)}{|\mathbf{r} - \mathbf{x}|}. \tag{120}$$

Here ξ is the transducer apodization function, and the vector $\mathbf{r}$ defines points on the surface of the transducer S. Note that for a single A-scan line, time is one dimensional - hence the use of scalar rather than vector time in the above two equations.

Spatial Sensitivity Functions and Spatial Resolution The spatial sensi-

tivity function (ssf) is defined as

$$ssf_{\mathbf{t}_o}(\mathbf{x}) = h(\mathbf{x}|\mathbf{t}_o); \tag{121}$$

and represents the phase-sensitive spatial weighting of the acoustic field to second-order variations in density and compressibility of the object at a snapshot of time. ssfs are important for computing statistical moments (such as covariance matrices) of speckle important for formal analysis of ultrasonic image quality. In Ref.[70] it was shown that the ssf rather than the psf is the desired quantity that defines the speckle covariance matrices K_+ and K_-. Ultimately it is these covariance matrices that the ideal observer uses to pre-whiten.

It will often be more convenient, however, to compute point-spread functions (psf), defined as the echo signals received from the transmitted pulse sequence probing a point scatterer at location $\mathbf{x} = \mathbf{x}_o$:

$$psf_{\mathbf{x}_o}(\mathbf{t}) = h(\mathbf{t}|\mathbf{x}_o) \tag{122}$$

, Spatial sensitivity functions and point spread funcitons have an important symmetry over isoplanetic patches, where local shift invariance may be assumed:

$$ssf_{\mathbf{t}_o=\mathbf{B}^{-1}\mathbf{x}_o}(\mathbf{x}) = psf_{\mathbf{x}_o}(-(\mathbf{t} - \mathbf{B}^{-1}\mathbf{x}_o)) \tag{123}$$

where $\mathbf{t} = \mathbf{B^{-1}x}$. This symmetry relation assumes that a 2D model of image formation is sufficient. (In 2D imaging, the ssf is a 3D function over space for a fixed point in time, and the point-spread function is a 2-D function over time for a fixed point in space). This symmetry breaks down for systems that are highly shift variant or when

a 2D object model is insufficient. A metric for assessing the size of an isoplanatic patch for shift-variant systems is discussed in.[?]

We are interested in computing the 2-D Fourier magnitude of the ssf in any given $x_1 - x_2$ plane for a given x_3, which we approximate using psfs:

$$\begin{aligned} |H(u_2, u_1)| &= |\mathcal{F}\{ssf_{\mathbf{t}_o=\mathbf{B}^{-1}\mathbf{x}_o}(x_2, x_1|x_3)\}| \\ &\cong |\mathcal{F}\{psf_{\mathbf{x}_o}(t_2, t_1)\}|_{t_2=2x_2/c, t_1=x_1/b} \end{aligned}$$

It is thus sufficient to calculate $h(x_2, t = t_1|x_1, x_3)$ and subsequently compute the Fourier transform along the x_2 and t dimensions to obtain from 119:

$$H(u_2, f|x_1, x_3) = i2\pi f H_y^2(f) V(f) \tag{124}$$

$$\times \quad [H_a(u_2, f|x_1, x_3) * H_a(u_2, f|x_1, x_3)] \tag{125}$$

where f is the temporal frequency, $H_a(u_2, f|x_1, x_3) = \mathcal{F}\{h_a(\mathbf{x}, \omega)\}$ and all other quantities are temporal Fourier transforms of the quantities in (119). The convolution is strictly over spatial frequencies u_2.

Fresnel Approximation

We begin our computation of $H_a(u_2, \omega|x_1, x_3)$ by using the Fresnel approximation[33] to compute

$$h_a(\mathbf{x}, f) \cong \frac{e^{ikx_1}}{i\lambda x_1} e^{i\frac{k}{2x_1}(x_2^2+x_3^2)} \int dS\xi(\mathbf{r}) e^{i\frac{k}{2x_1}|\mathbf{r}|^2} e^{-i\frac{k}{x_1}\mathbf{x}\cdot\mathbf{r}} \tag{126}$$

The Fresnel approximation is applicable for apertures that are weakly focused and paraxial points.

Lateral BW of Gaussian Apodized Transducer

In this section we aim to compute an analytic expression for the lateral bandwidth of spatial sensitivity functions of a Gaussian- apodized transducer. Without the assumption of Gaussian apodization, one obtains error functions with complex arguments and the analysis is less transparent. Let us consider a separable complex apodization function $\xi(r_2, r_3) = \xi_2(r_2)\xi_3(r_3)$ for quasi-planar transducers, where the azimuthal apodization functions can be considered a product of a real Gaussian apodization and a complex phase term representing focusing:

$$\xi_2(r_2) = e^{-\frac{r_2^2}{2\sigma_2^2}} e^{-jk[F-\sqrt{F^2+r_2^2}]} \tag{127}$$

where F is the focal length. For analytical convenience we do not impose any finite aperture - we simply assume that the Gaussian apodization is not severely truncated, i.e. that the aperture width L is significantly greater than σ_2. Additionally we find it advantageous to assume parabolic focusing by expanding the complex argument of (127) in a second-order Taylor series expansion in r_2 about 0 so that $F - \sqrt{F^2 + r_2^2} \cong r_2^2/(2F)$. With these approximations, Eq. (126) can be integrated by completing the square to become (see (153) and (154) of Appendix A):

$$h(x_2, \omega | x_1, x_3) \propto e^{-\Psi x_2^2} \tag{128}$$

with Ψ being a complex quantity $\Psi = \Psi_r + i\Psi_i$. The real part is given as

$$\begin{aligned}\Psi_r &= \left(\frac{k}{2x_1}\right)^2 \frac{1/2\sigma_2^2}{(\frac{1}{2\sigma_2^2})^2 + (\frac{k}{2})^2(\frac{1}{F} - \frac{1}{x_1})^2} \\ &\cong \frac{1}{2\sigma_2^2}\frac{F^2}{(F-x_1)^2} \end{aligned} \tag{129}$$

The approximation is true for points x_1 away from the focus when $x_1, F_2, \sigma_2 \gg \lambda$. Similarly,

$$\begin{aligned}\Psi_i &= \left(\frac{k}{2x_1}\right)^2 \frac{(\frac{k}{2})(\frac{1}{F} - \frac{1}{x_1})}{(\frac{1}{2\sigma_2^2})^2 + (\frac{k}{2})^2(\frac{1}{F} - \frac{1}{x_1})^2} + \frac{k}{2x_1} \\ &\cong \left(\frac{k}{2}\right)\left(\frac{1}{x_1 - F}\right) \end{aligned} \tag{130}$$

The real part defines a lateral Gaussian envelope $e^{-\Psi_r x_2^2}$ for h_a, and the imaginary part defines a linear spatial frequency phase modulation $e^{-i(\Psi_i x_2)x_2}$, i.e. a baseband chirp. We need to transform our result to the spatial frequency domain to compute the lateral bandwidth:

$$H_a(u_2, \omega) \propto \mathcal{F}\{e^{-\Psi x_2^2}\} \propto e^{-\Sigma u_2^2} \tag{131}$$

where

$$Re\{\Sigma\} = \pi^2 \frac{\Psi_r}{\Psi_r^2 + \Psi_i^2} \cong \pi^2 \frac{2F^2}{\sigma_2^2 k^2} \tag{132}$$

Eq. 131 can be substituted into the expression for $|H|$, Eq. 124 to get that

$$|H(u_2, \omega | x_1, x_3)| \propto e^{-Re(\Sigma)u_2^2} \tag{133}$$

which means that the lateral -6 dB BW for all axial locations x_1 (that satisfy the approximations of the model) can be written as

$$BW_{lat} = 2.35\frac{2}{\sqrt{Re(\Sigma)}} = 2.35\frac{\sigma_2 k}{F} \tag{134}$$

which is a constant for all axial depths, and equal to the reciprocal of the expected focal resolution! This is a rather remarkable result which we validate using FIELDII simulations in the next section.

It is illuminating to remind ourselves that the temporal bandwidth of a pulse propagating in a medium remains relatively unchanged over time. Interestingly the same principle holds for the *lateral* bandwidth.

8.3.2. FIELD II Validation of Constant BW Prediction

We simulated Gaussian apodized transducers using FIELD II, and computed the 2-D Fourier magnitude. The lateral spatial frequency profile at the center frequency was normalized by its total area so as to mimic a Gaussian probability distribution. This was then integrated against the second moment of spatial frequency u_2^2 to estimate the Gaussian sigma parameter, which is our measure of lateral BW. The results of the simulation validate the constant BW prediction of the previous section.

This result means that we may obtain spatial resolution comparable to the focal region but throughout the entire field of view. Dynamic focusing and aperture growth techniques can also perform similar functions, but do not recover all available energy, as will be discussed shortly.

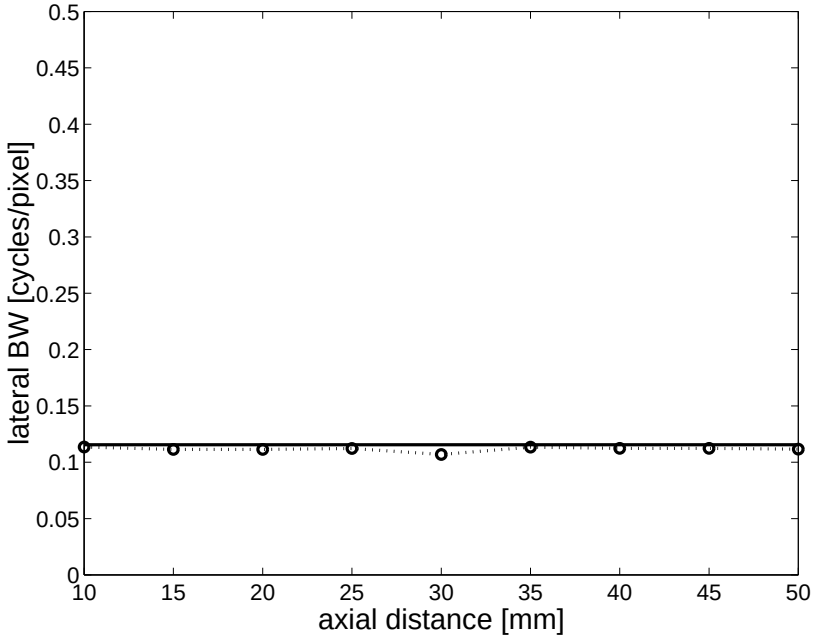

Figure 35. Lateral BW of Gaussian Apodized Transducer as a function of axial depth. Parameters to be added.

8.3.3. Extension of the concept of time-BW product

The product of the time-duration of a coded waveform with its bandwidth is termed the time-bandwidth product (TBP). The TBP is a unitless quantity representative of potential information that is one for simple pulses and greater than one for coded waveforms. It is appropriate to extend the TBP concept to spatial coding. Here we define a quantity which we shall call the Lateral Space-BW product (SBP_{Lat}) which is given as the product of the lateral spatial extent of the psf times the lateral BW of the psf. For the Gaussian apodized transducer that we have been analyzing,

$$SBP_{Lat} = BW_{Lat}\sigma_{Lat} \cong 1 + (\frac{k\sigma_2^2|F - x_1|}{F^2})^2 \tag{135}$$

where σ_{Lat} is the -6 dB width of the spatial lateral extent of the Gaussian aperture apodization. As might be expected, the SBP_{Lat} is one at the focus (no wavefront curvature thus no lateral coding). It is greater than 1 away from the focus and is greater for distances far from the focus as long as the Fresnel approximation is obeyed.

8.3.4. Energy Content

Often the TBP for axial codes is associated with recoverable acoustic energy. Two codes can be compared for energy content, however, only when waveform amplitudes are identical. For spatial codes, although wavefront curvature is the principle determinant of the SBP_{Lat}, it should be noted that there are losses in energy with increasing depth - both diffractive losses and losses due to attenuation. Diffractive losses are due to energy that radiates away from the main beam and are not recoverable. The amount of recoverable energy is determined by the solid angle subtended between the field point and the detector edges. Spherically scattered waves from the field point will only be recorded if they are backscattered to points on the detector surface. Diffractive losses explain why the ideal observer SNR_I in Fig. 32 falls off with increasing depth. The combination of good spatial resolution and good echo SNR explain why the nearfield is predicted to be more informative than the focal region or the farfield. Attenuation is another very important source of loss but it was not accounted for in Fig. 32. Attenuation can easily be included into the analysis in future work.

8.3.5. Spatial Filtering for Swept Point Source Transducers

It will be instructive to now consider a point source scanned laterally (transmission and reception at each lateral location). Small array elements behave effectively as point sources in the farfield of the element. Consideration of the point-source problem will help us understand how our technique compares with Synthetic Receive Aperture (SRA) and Dynamic Receive Focusing (DRF) methods. Additionally offer a new and important interpretation to spatial matched filtering methods as a delay-and-sum image reconstruction technique. Eventually we will return to the problem of imaging in the nearfield of a focused transducer with large aperture (in contrast to a point source as we now investigate).

In addition to the assumption of a point source, let us consider a delta-function transmission, and an ideal electro-mechanical coupling. Then, neglecting a cosine directionality term (for simplicity)

$$h(x_2, t|x_1) = \frac{1}{4\pi^2} \frac{\delta(t - 2\sqrt{x_1^2 + x_2^2}/c)}{x_1^2 + x_2^2} \tag{136}$$

The ideal observer test statistic tells us that the reconstruction strategy of choice is spatial matched filtering. The corresponding spatial matched filter is thus given as $f_r(x_2, t) \equiv h(-x_2, -t|x_1)$. RF data g can be reconstructed to form the reconstructed signal y, given as $y(x_2, t) = h(-x_2, -t|x_1) * g(x_2, t)$. Note that our choice of a matched filter here lacks a time convolution with the excitation pulse. This is intentional. Including the added pulse convolution in the matched filter will have the effect of broadening axial speckle. With Wiener filtering, some of this broadening can be

recovered. Instead, however, we currently desire to present a technique of spatial matched filtering that does not broaden axial speckle.

Denoting $\frac{1}{4\pi^2}\frac{1}{x_1^2+x_2^2}$ as $w(x)$, we can write y as

$$\begin{aligned} y(x_2,t|x_1) &= \int\int f_r(x_2-x_2',t-t')g(x_2',t')dx_2'dt' \\ &= \int\int h(x_2'-x_2,t'-t)g(x_2',t')dx_2'dt' \\ &= \int\int w(x_2'-x_2|x_1)\delta(t'-t-2\sqrt{x_2^2+(x_2'-x_2)^2}/c)g(x_2',t')dx_2'dt' \\ &= \int w(x_2'-x_2|x_1)g(x_2',t+2\sqrt{x_1^2+(x_2'-x_2)^2}/c)dx_2' \end{aligned} \tag{137}$$

The limits of integration of x_2' are over some receive aperture width (there could also be an additional receive apodization $\xi(x_2')$ applied to the weighting function w). The last form means that the spatial matched filtering reconstruction operation can be viewed as a delay and sum procedure as shown in Fig. 36 with delay

$$\tau_{SF}(x_1,x_2) = -2\sqrt{x_1^2+x_2^2}/c \tag{138}$$

evaluated at lateral location $x_2'-x_2$.

This important observation will eventually be extended to spatial coding with curved wavefronts in the nearfield of larger focused apertures. Before doing so we wish to compare and contrast our method with that of synthetic receive aperture (SRA) focusing methods. SRA is the natural technique to compare with because like the swept point-source situation we have been modeling, SRA imaging also uses small aperture transducers scanned to transmit and receive at different locations over time to synthesize an effectively larger aperture (See Fig. 37).

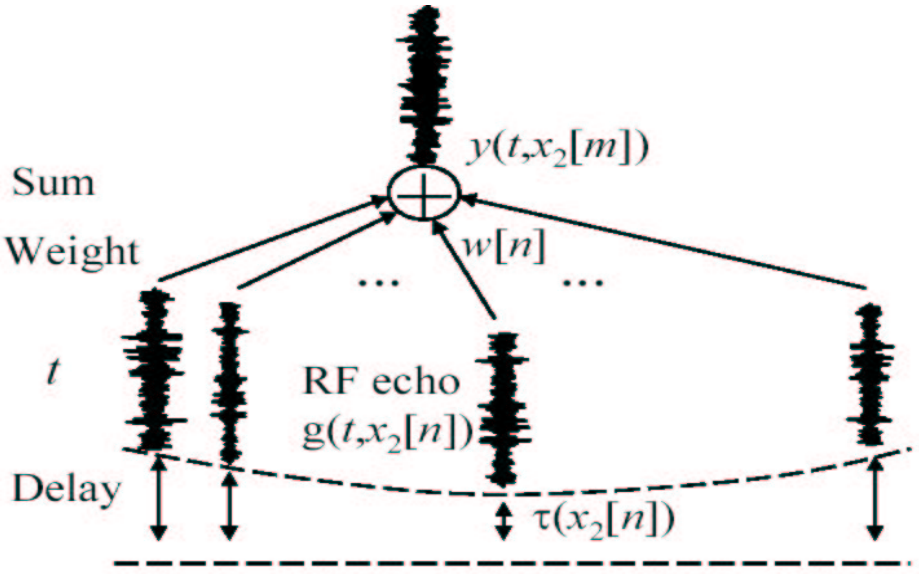

Figure 36. Spatial matched filtering can be viewed as a delay and sum operation.

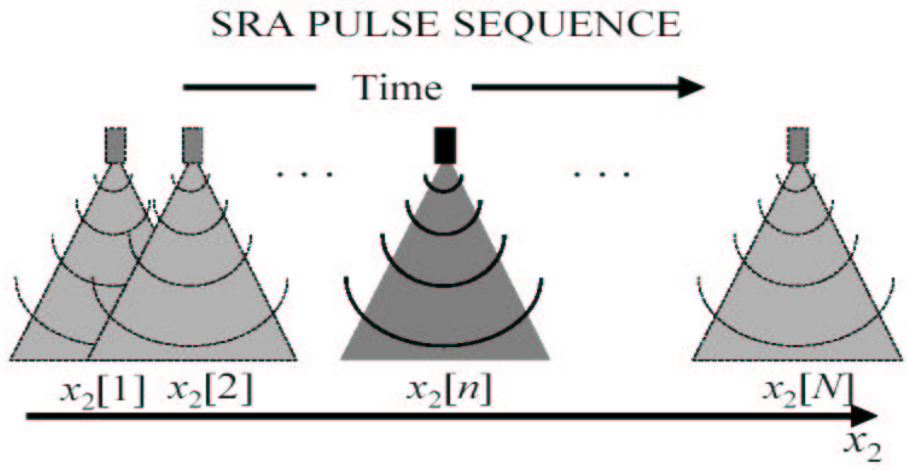

Figure 37. The SRA pulse sequence is identical to the one we consider for spatial focusing in this section. Here the curved line represent transmission wavefronts.

8.3.6. Comparison with Synthetic Receive Aperture Methods

SRA and spatial filtering methods are similar in that they delay and sum adjacent A-scan lines with weighting functions to reconstruct an image. There are important

differences, however, as will shortly be seen.

SRA methods have been used recently for multi-row arrays to transmit and receive one row at a time - thus using the same channel count as standard linear array systems.[22] In this situation an elevational column of elements is used to synthetically focus at each receive depth similar to dynamic receive focusing.

Here we consider reconstruction in the lateral direction. For simplicity, consider the reconstructed A-scan line at lateral position $x_2 = 0$, we would thus delay and sum RF element acquitions $g(x_2[n], t)$ for element positions $n = 1, ..., N$ as follows:

$$y_{SRA}(x_2 = 0, t) = \sum_{n=1}^{N} w(x_1, x_2[n]) g(x_2[n], t - \tau_{SRA}(x_1, x_2[n])) \tag{139}$$

where w is a synthetic receive apodization function, and the delay is given by geometrical arguments as

$$\tau_{SRA}(x_1, x_2[n]) = x_1/c - \sqrt{x_1^2 + x_2^2}/c. \tag{140}$$

This delay is meant to simulate receive focusing at each depth. No focusing is done on transmission. The effective pulse-echo psf is thus a convolution between an unfocused transmit field and a focused receive field. Consequently, only a fraction of the transmitted energy is recovered. In fact, SRA focusing can be viewed as discarding energy away from the receive focus, as illustrated in Fig. 38.

Our approach is different. Notably, τ_{SF} differs from τ_{SRA} by a multiplicative factor of 2 and an additive factor of x_1/c. The additive factor of x_1/c is of little significance as it simply shifts the reconstructed signal in time. The factor of 2 is significant.

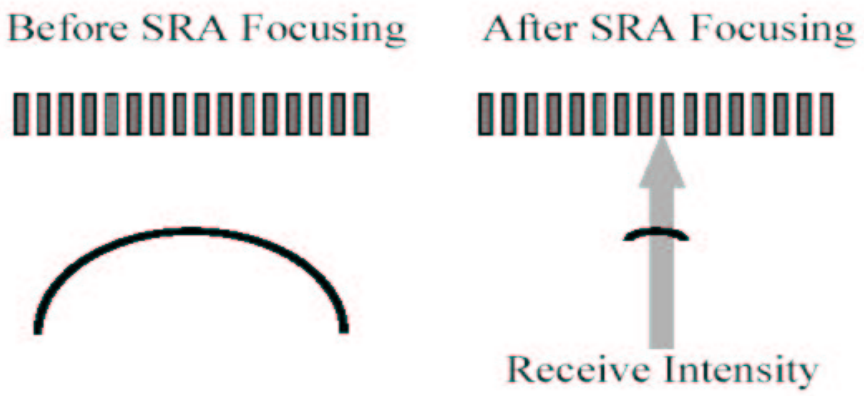

Figure 38. Here the curve represents the pulse-echo psf. SRA Focusing recovers only a fraction of the energy in the psf.

Rather than focusing at each receive position, our method aims to restore coherence, recover spatial resolution, and recover energy. Eq. 147 can be visualized graphically as in Fig. 39. The curved wavefronts represent a pulse-echo point-spread function **g** due to a 3-cycle transmission excitation. To form a single A-scan of the reconstructed image, each point of the wavefront (each A-scan) is shifted with a position dependent time-delay. A lateral sum is then performed. At the appropriate axial and lateral position, the wavefront is flattened, and the lateral sum tends to compress all the energy along the wavefront into a small volume. Away from the reconstruction maximum, the delay operation re-shapes but does not flatten the wavefronts - and the lateral summation coherently integrates to a small (and hopefully negligible) value.

Besides recovering much of the transmitted energy and restoring coherence by flattening the wavefronts, spatial filtering aims to recover spatial resolution intrinsi-

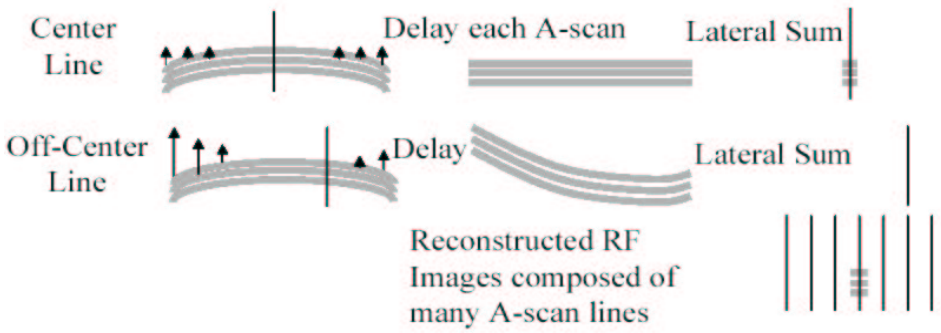

Figure 39. Graphical representation of Eq. 147

cally phase-coded in the wavefront curvature. In fact, we argue that SRA methods do not fully recover spatial resolution encoded in the wavefronts. Evidence of this is that coherence is not fully restored: SRA often demonstrate spatially compact psfs that exhibit curved wavefronts. The leftover wavefront curvature represents lateral spatial resolution that is not fully recovered. Instead of showing data to this end now, we choose to include experimental data for the similar and more common method of dynamic receive processing. This will be done in the experimental results section.

Thus far we have largely discussed spatial matched filtering. The ideal observer strategy also tells us that Wiener regularization may be advantageous as it can whiten remaining RF correlations. Spatial matched filtering works well alone without Wiener whitening when matched filtered data samples are nearly independent. The degree to which additional whitening is successful is limited by the amount of electronic noise, and our ability to accurately model the whitening filter.

8.3.7. Spatial Resolution Limitations for Swept Point Source Imaging

Soon we will extend the delay and sum interpretation of spatial matched filtering to the nearfield of focused transducers. Before doing so, it is important to consider the fundamental limits of spatial resolution for the case of laterally scanned point source excitations. Specifically, we consider the resolution limits of spatial matched filtered psfs without Wiener regularization, or, equivalently, in noisy environments when Wiener filtering and matched filtering give the same results. Additionally, we consider impulse temporal excitations and ideal electro-mechanical transducer responses.

Reconstruction of noiseless RF data from an impulse-excited psf due to object function $f(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_o)$ at depth x_1 by spatial matched filtering is thus given as $h(x_2, t) * h(-x_2, -t)$ which we write as:

$$\begin{aligned} R_{hh}(x_2, t|x_1) &= \left(\frac{1}{4\pi^2}\right)^2 \int\int \frac{\delta(t' - 2\sqrt{x_1^2 + {x_2'}^2}/c)}{x_1^2 + {x_2'}^2} \\ &\times \frac{\delta((t - t') - 2\sqrt{x_1^2 + (x_2' - x_2)^2}/c)}{x_1^2 + (x_2' - x_2)^2} dx_2' dt \\ &= \left(\frac{1}{4\pi^2}\right)^2 \int \frac{\delta(t - \Delta\tau)}{(x_1^2 + {x_2'}^2)(x_1^2 + (x_2' - x_2)^2)} dx_2' dt \end{aligned} \tag{141}$$

where

$$\Delta\tau = 2\sqrt{x_1^2 + (x_2' - x_2)^2}/c - 2\sqrt{x_1^2 + {x_2'}^2}/c. \tag{142}$$

In this derivation we have used the identity $\delta(x) = \delta(-x)$.

R_{hh}, is an important quantity that determines the spatial resolution of the system as is its Fourier transform $|H(u_2, f)|^2$. This is in turn fundamental to the ideal observer performance, Eq. 115. With receive apodization $\xi(x_2')$ we in fact are doing

mismatched filtering rather than matched filtering - thus the autocorrelation reduces to the cross correlation

$$R_{hh'}(x_2,t|x_1) = \int \xi(x_2') \frac{\delta(q(x_2'))}{(x_1^2+x_2'^2)(x_1^2+(x_2'-x_2)^2)} dx_2' \tag{143}$$

where $q(x) = 2\sqrt{x_1^2+x^2}/c - t - 2\sqrt{x_1^2+(x-x_2)^2}/c$. Unfortunately the Fresnel approximation is not entirely appropriate for point sources since the resulting diverging spherical waves are not paraxial beams. The equation can nevertheless be solved analytically using the relationship $\delta(q(x)) = \sum_i \frac{1}{|q'(x_i)|}\delta(x-x_i)$ where x_i are roots of $q(x)$ and the derivatives are such that $q'(x_i) \neq 0$.[23] $q(x) = 0$ reduces to a quadratic equation $\alpha x^2 + \beta x + \gamma = 0$ with

$$\begin{aligned}
\alpha &= -4x_2^2 - (tc/2)^2 \\
\beta &= 4x_2[-4x_1^2 - x_2^2 + (tc/2)^2] \\
\gamma &= -4x_1^2(tc/2)^2 - 2(tc/2)^2 x_2^2 + x_2^4 + (tc/2)^4
\end{aligned}$$

The roots, given by the quadratic formula, are functions of t and x_2 for a given axial depth x_1. The resulting $R_{hh'}$ can be written as:

$$R_{hh'}(x_2,t|x_1) = \sum_i \frac{1}{|q'(x_i)|} \frac{\xi(x_i)}{(x_1^2+x_i^2)(x_1^2+(x_i-x_2)^2)} \tag{144}$$

where the derivative is given as

$$q'(x) = \frac{2}{c}\left(\frac{x}{\sqrt{x_1^2+x^2}} - \frac{x-x_2}{\sqrt{x_1^2+(x-x_2)^2}}\right). \tag{145}$$

The important point here is that the ideal autocorrelation of the psf, R_{hh}, is not a simple delta function - even for delta function excitations. (Note however that

144 singularities). R_{hh} describes the fundamental resolution limit of the technique. Transmit waveform shape and finite transmit aperture width will add additional blurring to $R_{hh'}$. Notice that during the entire discussion, we have assumed continuous sampling. Meaningful analysis and discussion of sampling effects may arise from a discrete analysis which we forego for sake of conciseness.

So far we have only discussed synthetic receive processing. There are also synthetic transmit- and synthetic transmit-receive approaches which require a 'complete set' of data. By this we mean that one transmits at one lateral location and receives at another, for an entire range of N transmit and N receive locations (N^2 echo signals). This allows the capability to focus synthetically on transmit as well as receive. We expect similar spatial filter analysis for this higher-dimensional problem.

8.3.8. A depth-varying delay and sum beamforming approach to synthetic aperture reconstruction of a large aperture focused transducer

Synthetic receive techniques that use small elements to approximate point sources can recover less energy per scan line than a larger aperture scanning the identical lateral distance. Here we present a method for synthetic aperture reconstruction for a larger aperture with fixed transmit-receive focal length. The hope is to be able to improve the SNR over small aperture synthetic imaging approaches. A competing approach is to use virtual sources as described in Ref.[85] - an approach which also uses unfocused beams and performs reconstructions - and thus may be analyzed using our framework. Our approach unlike others may be very suitable for building a single element

(mechanically scanned) system with depth of field and spatial resolution comparable with a much more expensive array system. This approach is particularly attractive for very high frequency imaging where array technology is still under development. Similarly, it could be a means of reducing the channel count (but not necessarily the number of elements) and hence the cost of current array systems. These ideas will be explored more fully in subsequent sections.

Our strategy for this is to use the computations done previously for a Gaussian-apodized transducer, deduce the matched filter, and estimate the depth-dependent delays needed to compress the energy in a large curved wavefront into a compact volume. It will be seen that edge waves due to aperture truncation will be a confounding factor for reconstruction. Apodization is a means of abating the effects of these artifacts.

Taking the inverse Fourier transform of Eq. (128) we have that (Eq. 167 of Appendix A)

$$h_{impulse}(\mathbf{x}, t) = p(\mathbf{x}, t) * \delta(t - \tau_F(x)) \tag{146}$$

where $*$ represents a temporal convolution and $p(\mathbf{x}, t)$ is given by 168 of Appendix A. For pulse-echo imaging, the appropriate delay for synthetic receive processing is thus τ_F as given in Appendix A. This delay is important because it can allow us to re-write the spatial matched filter for large aperture focused transducers as a delay and sum procedure. Similar to Eq. (147) the reconstructed RF signal can be written

as

$$y(x_2, t|x_1) = \int g(x_2 - x_2', t - \tau_F(x_1, x_2 - x_2')|x_1)dx_2' \tag{147}$$

The procedure can be visualized by considering the RF image of a single point scatterer - a curved wavefront. Synthetic aperture processing for a given A-scan line applies delays to all A-scan lines such that the wavefront is flattened (almost - but not perfectly, due to a quadratic rather than linear arguments of the delta functions in 141). Then the result is integrated laterally to give the desired post-processed A-scan line. In contrast to dynamic receive processing the proposed method attempts to concentrate all the energy along the wavefront into a point rather than focus on a segment of the wavefront (this eliminates energy away from the receive focal region). From this perspective it is easy to see that any energy off of the curved wavefront (for example - the X-like pattern due to edge waves in Fig. 33(a) will integrate to give undesirable artifacts as was the case with matched filtering.

Two approaches exist for implementation. (1) Construct matched filters and implement the delay and sum operations by fast 2-D FFT convolution operations. A collection of depth dependent filter banks could be used to accomplish this over isoplanatic patches. (2) The depth-dependent delay and sum procedure can be implemented with interpolation routines to avoid unnecessary time quantization errors. It should be emphasized that the synthetic aperture delay and sum operation is applied to the beamformed RF image data g, and not the pre-beamformed echo signals. Envelope detection can subsequently be performed on y.

8.3.9. Potential advantages of spatial coding over dynamic receive and synthetic receive processing

As seen in Fig. 40, dynamic receive processing recovers only a fraction of energy in transmitted wavefronts. Although not illustrated in the figure it is important to note that there is a potential advantage in spatial resolution over dynamic receive techniques - something that will be demonstrated experimentally in a later section.

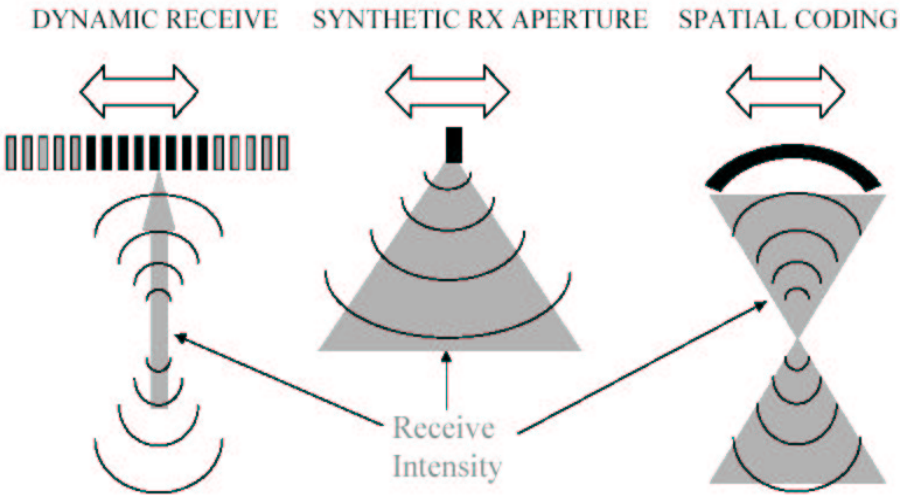

Figure 40. Comparison of different imaging techniques.

8.4. Experiment. Performance of Unfocused Imaging on the Siemens Antares Ultrasound System

We used a programable Siemens ANTARES ultrasound scanner to test some of our ideas. This scanner possessed an ultrasound research interface (URI) that allowed us to control acquisition parameters not accessible in clinical mode, and to save RF to files for offline analysis. A library of MATLAB functions (offline processing tool or

OPT) for reading and processing the data was available to us to assist in our analysis. Also available was a MATLAB package to simulate point-spread functions and images with Antares parameters read from probe and system header files.

8.4.1. Nearfield psfs

To measure the nearfield psf of a fixed-focus beam on the ANTARES, we used the URI to turn off dynamic receive, aperture growth and receive apodization functions. We set the receive f-number to 2.1, the transmit focus at 4cm and the receive focus at 3.9cm (the URI allows only several values for these parameters). To image psfs we simply aquired RF data from sparse microparticulate matter in water at low transmit amplitudes. The measured and corresponding simulated psf is shown in Fig. 41. The curved wavefront of the simulated psf is similar to the measurement, however, there is significantly more edge wave energy experimentally. Proprietary beamforming techniques used on the system were the suspected cause. This will turn out to be a confounding source of contrast reduction in implementing our unfocused imaging techniques.

8.4.2. Low Contrast Lesion Phantoms: Experiment and Simulations

To experimentally test the methods of unfocused imaging, we used the Siemens Antares system with a VF10-5 linear array transducer to scan graphite-agar phantoms with cylindrical channels. Size-matched hyper-echoic rods of graphite-agar material could be inserted into the channels and rotated or translated to give independent speckle realizations inside the inclusion. We used channels with radii of 1.5 and 3

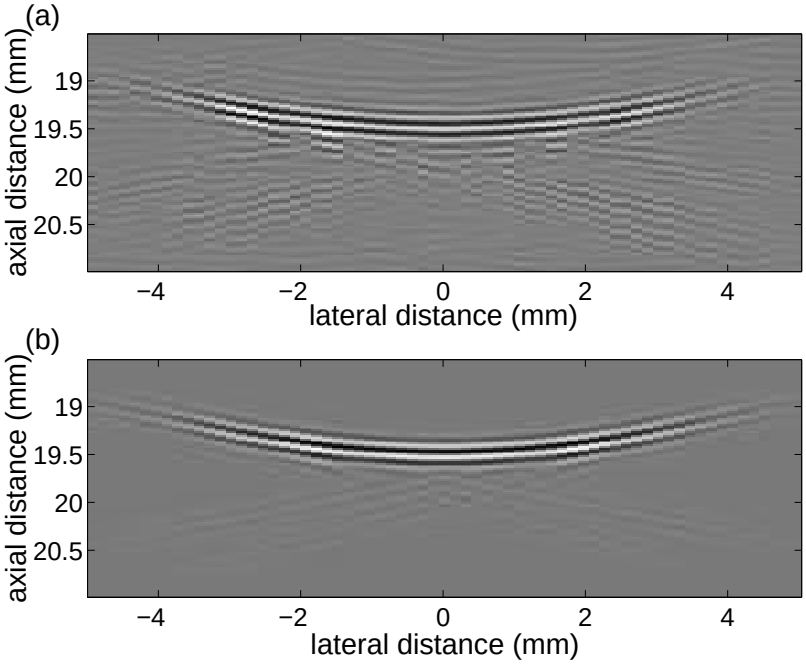

Figure 41. (a) measured nearfield psf (b) simulated psf. Parameters to be added.

mm respectively. We used a transmission frequency of 6.67 MHz for all experiments. For both nearfield and focal imaging, we chose an f-number of 2.1, and chose the transmit and receive focal distances as 4 and 3.9 cm, respectively. For nearfield imaging, the inclusion centers were positioned at approximately 2 cm in depth. For focal region imaging, the inclusion centers were at 4 cm.

In this section we report image quality measurements from low contrast hyperechoic inclusions. The purpose of this section is to compare C, CNR, and S_c in (1) nearfield B-mode processed images (2) nearfield matched filtered images (3) focal region B-mode images. Sample images are shown in Fig. 42. Image quality metrics, reported in Table 2, reflect 10 independent scans. We chose to use matched filtering rather than Wiener filtering because the measured psf was too noisy to use

successfully in a Wiener filter, and we could not have full knowledge of the system operations to simulate the psfs exactly. Future improvements in modeling or measurement techniques should allow Wiener filtering to be implemented with better results.

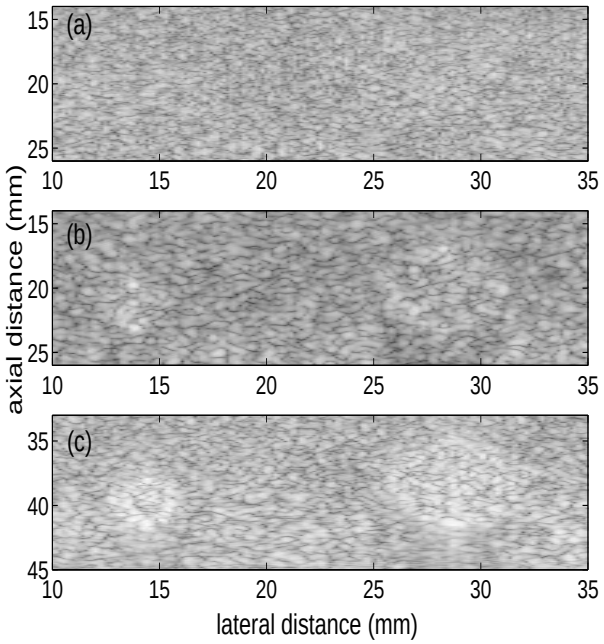

Figure 42. (a) Nearfield, unfiltered (b) Nearfield, filtered, (c) focal B-mode. Focal distance was 4 cm, F-number was 2.1. VF10-5 transducer parameters. The grayscales in all images were normalized to the same standard deviation.

Note that the filtered speckle length is broadened! This is consistent with the restoration of phase coherence (in our situation this means the re-flattening of wavefronts). A related topic concerns speckle texture in the presence of phase aberration:[86] found that speckle was broadest when aberrations were least - again a manifestation of coherence.

Table 2. Low Contrast Hyperechoic

Big Lesion

Region	Method	C	CNR
near	BM	0.10 ± 0.02	0.17 ± 0.03
near	MF	0.32 ± 0.02	0.46 ± 0.03
focal	BM	1.17 ± 0.05	1.37 ± 0.05

Small Lesion

Region	Method	C	CNR
near	BM	0.15 ± 0.05	0.26 ± 0.09
near	MF	0.58 ± 0.07	0.77 ± 0.08
focal	BM	1.77 ± 0.09	1.9 ± 0.1

Lateral Speckle Lengths

Region	Method	S_c
Near	BM	1.5 ± 0.1
near	MF	2.2 ± 0.2
focal	BM	2.2 ± 0.3

8.4.3. Anechoic Lesion Phantoms: Experiment

Using identical scanning parameters to the previous section we next report image quality measurements from anechoic channels. Again, we compare C, CNR, and S_c in (1) nearfield B-mode processed images (2) nearfield matched filtered images (3) focal region B-mode images. The results, shown in Table 3, reflect measurements of the image shown in Fig. 43. It can be seen that the images of the inclusions in the nearfield are much improved with filtering, but inferior to the focal region. In investigating high contrast anechoic channels, we are in fact no longer considering low contrast detection tasks, but rather the task of discriminating between an anechoic cyst and a solid mass, therefore the ideal observer strategy in Eq. (3) may not apply. The nearfield is clearly inferior to the focal zone for this task, due to spatial sidelobe artifacts in the channel interior. To be effective for mass-cyst discrimination tasks, spatial sidelobe levels need to be reduced considerably.

8.4.4. Testing the 2-D Linear Systems Model of Image Formation: Theory v.s. Experiment

As a test that the linear systems theory could accurately model nearfield speckle texture and mimic expected image quality, we compared experimental and simulated phantom images, and their corresponding image quality metrics. Images of pre- and post- filtered images are shown in Fig. 44 and quality metrics for the small lesion are given in Table 4. Very reasonable agreement was obtained, suggesting that the 2-D speckle model of image formation (on which the detection theory is based) could

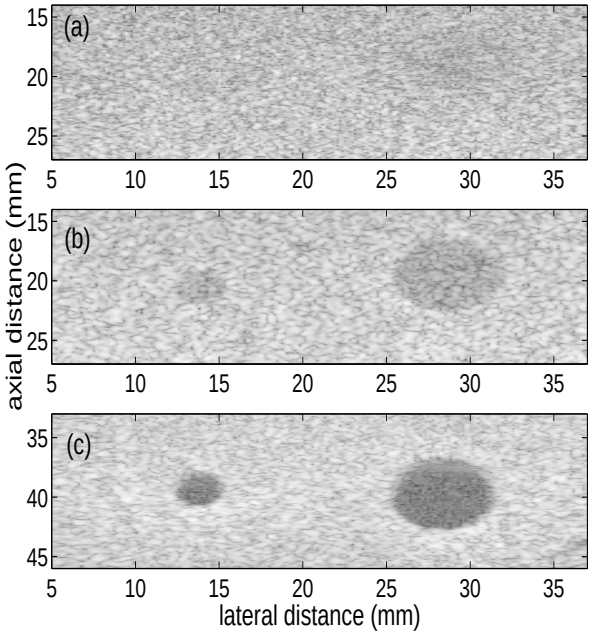

Figure 43. (a) Nearfield, unfiltered (b) Nearfield, filtered, (c) focal B-mode. Parameters to be added.

actually reproduce experimental images.

8.4.5. Explaining Performance

Just because the experimental unfocused results are not as high quality as focused images does not mean that our theoretical framework is not valid. We give two significant reasons for this.

No Whitening

First, we have been using matched filtering rather than Wiener filtering techniques. Matched filtering approaches do not attempt to whiten the resulting speckle correlations in the filtered image as the Wiener filter can. We have not implemented

Table 3. Anechoic

Big Lesion

Region	Method	C	CNR
near	BM	-0.22	0.42
near	MF	-0.63	1.25
focal	BM	-0.89	2.0

Small Lesion

Region	Method	C	CNR
near	BM	-0.05	0.10
near	MF	-0.56	1.10
focal	BM	$*$	$*$

Table 4. Image Quality of Simulations vs Experiments

Small Lesion

Data	Method	C	CNR	S_c
Sim	BM	0.04	0.08	1.7 ± 0.1
Expt	BM	0.05	0.10	1.5 ± 0.1
Sim	MF	0.52	1.2	2.1 ± 0.1
Expt	MF	0.56	1.1	2.2 ± 0.2

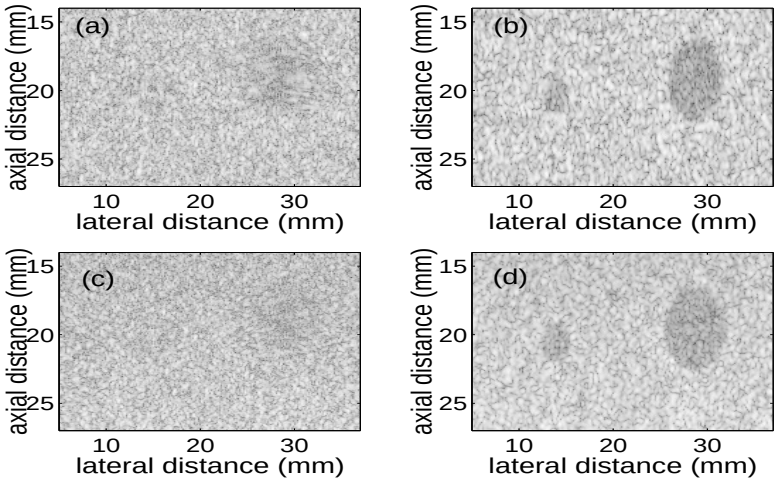

Figure 44. (a) Simulation: Nearfield, unfiltered (b) Simulation: Nearfield, filtered, (c) Experiment: Nearfield, unfiltered (d) Experiment: Nearfield filtered. Parameters to be added.

Wiener filtering because we do not have full knowledge of the ANTARES beamforming operations to simulate the psfs well enough to implement an effective Wiener filter. This brings us to the next point.

Edge Wave Energy

The ANTARES psfs had more edge wave energy than the simulations in due to default proprietary beamforming techniques. Edge waves are a source of contrast degradation for our technique because they represent energy away from the main wavefront and cannot be adequately suppressed by matched filtering (Wiener filtering may do better at this since edge waves are a source of correlation which the Wiener filter may be able to partially whiten). Should it be possible to implement beamforming techniques that give measured psfs similar to the simulations in the

Theory section (minimizing edge wave energy), much improved experimental performance may be expected.

8.4.6. Reduction in Edge Wave Energy

To test this hypothesis, we disabled the default beamforming sequence responsible for the excessive edge wave energy. The resulting psf is shown in Fig. 45(a) below. A reduction of edge wave energy compared to Fig. (41) is evident.

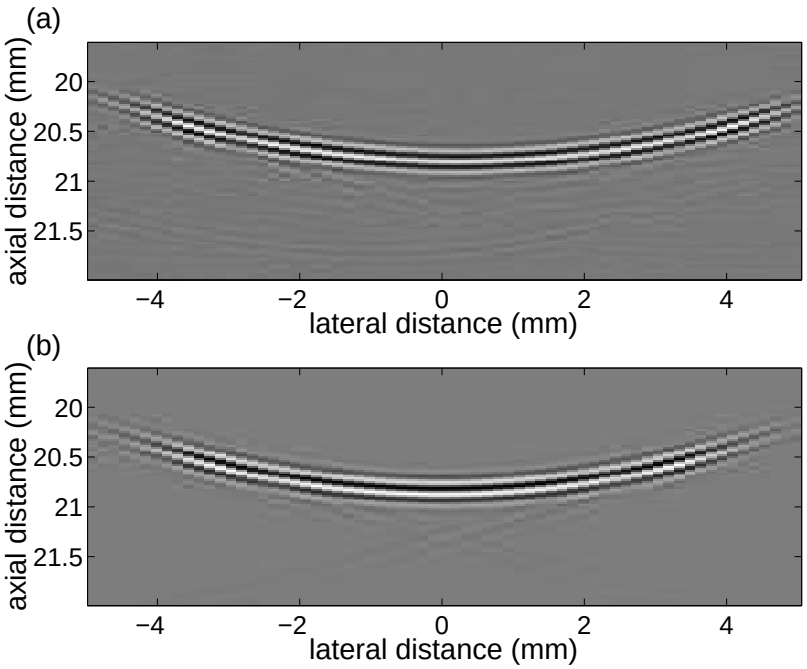

Figure 45. Proprietary beamforming option disabled: Note the reduction of edge wave energy compared to Fig. (41). (a) measured nearfield psf (b) simulated psf. Parameters to be added.

Image reconstruction performance is much better when edge wave energy is reduced, as can be seen by comparing Fig. 46(a) with Fig/ 43(b). Additional image

quality comparisons with unfocused unfiltered images and focal region images can be made by comparing with Tables 3 and 2.

Additional reduction in edge wave energy may be imposed by aperture apodization, however, at the cost of spatial resolution. More research is needed to understand the best apodization schemes for unfocused imaging.

8.4.7. Comparison with Dynamic Receive Focusing

Spatial coding with curved wavefronts and matched filtering reconstruction is seen to have better echo SNR, better contrast and CNR compared with dynamic receive focusing, as can be seen in Fig. 46 and Table 5 . Both lesions were imaged at a depth of 2 cm, and in both cases the transmit focus was set at 4 cm. The F-number in (a) was 2.1. It should be qualified that images were acquired at low power (10% of the default power level) to demonstrate the SNR advantage. The experimental results are consistent with predictions in Section 3.H. As illustrated in Fig. 40 dynamic receive focusing methods discard energy in curved wavefronts, whereas the our method attempts to recover all the energy in the wavefront, as illustrated in Fig. 39. Note the improvement in the spatial resolution between Figs. 46(a) and (b). It should be qualified that proprietary aperture growth techniques were enabled along with dynamic receive focusing, and thus we are not sure what the instantaneous F-number was for Fig. (b). It may be that (a) used a wider receive aperture. Although this could be one factor contributing to the inferior spatial resolution in (b) compared to (a), it should be emphasized that the dynamic receive focusing technique does not

restore coherence (here coherence is directly related to the flatness of the wavefronts) as does spatial matched filtering. This is apparent by observing that speckle in (b) is much finer than in (a). Again we emphasize that speckle is broadest when coherence is greatest. Coherence ensures that energy is concentrated locally rather than spread out in complex constructive and destructive interference patterns.

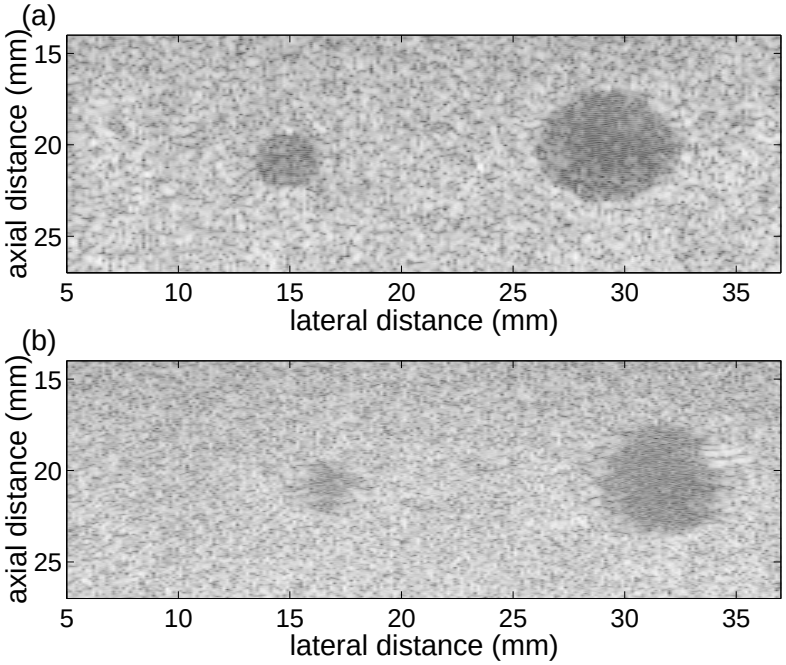

Figure 46. Anechoic Lesion Phantom: (a) Nearfield filtered image (default proprietary beamforming option disabled). F-number=2.1. (b) Dynamic receive focusing image. In both images the transmit focus is at 4 cm.

8.5. Discussion

A fundamental figure of merit for detection tasks, the ideal observer SNR, was computed over the depth of a focused beam to obtain Fig. 115. SNR_I is a measure of

Table 5. Anechoic Lesions: Comparison of Nearfield Imaging (Default Beamformation Disabled) with Dynamic Receive Focusing

Big Lesion

Region	Method	C	CNR
near	MF	0.75	1.5
near	DRF	0.61	1.0

Small Lesion

Region	Method	C	CNR
near	MF	0.71	1.4
near	DRF	0.45	0.8

the statistical separation of signal present and signal absent test statistic distributions and parameterizes the receiver operating characteristic curve. It is a measure of the diagnostic information available in the unprocessed RF data with a given system design. It is useful because it may help us understand how to maximize diagnostic information content. We have developed a set of analytical insights that can help us interpret Fig. 115. Similar to the constant nature of axial bandwidth over propagation, it was seen that the lateral bandwidth of a pulse remains approximately unchanged as it propagates. The curved wavefronts at points away from the focus, however, are spatially distributed, giving a greater Space-Bandwidth product away from the focal region. The amount of energy in a curved wavefront, however decreases away from the transducer due to diffractive and other losses. Detection performance relies on both energy and spatial resolution. Performance is expected to be best in the

nearfield, where there is comparable resolution to the focal zone but more energy. Performance decreases with energy losses away from the transducer. The ideal observer prediction motivated closer investigation into nearfield imaging techniques. The form of the ideal observer test statistic also motivated spatial matched and Wiener filtering techniques. For a given system design, the ideal observer attempts to match filter then pre-whiten remaining correlations. The ideal observer SNR then allows us to investigate what designs are the most informative. This paper focuses on one possible design technique. The ideal observer approach could be used to investigate many other designs.

When we tested the unfocused imaging technique with the ANTARES system (with the proprietary beamformation functions as default) - we found that the image quality metrics were superior for unfocused filtered images compared with unfocused unfiltered images but inferior to focused images. Note however that the ANTARES system psfs were different that the hypothetical designs that we simulated. In particular, the ANTARES psfs had more edge wave energy than the simulations. When we turned off these functions, edge wave energy was reduced significantly, and reconstructed image quality improved considerably.

Future work could aim to simulate system psfs well enough to effectively implement Wiener filtering and thus reduce undesirable correlations (through speckle whitening) even further.

It is significant that we have demonstrated an improvement over dynamic receive

processing since this is the technique most widely used on current array-based ultrasound imaging systems. The important point is that spatial matched or Wiener filtering can restore coherence by flattening and concentrating energy in the curved wavefronts similar to synthetic receive aperture techniques, thus improving spatial resolution and echo SNR, and thus also contrast and CNR.

We might ask what would happen if we were to set the transmit focus in the region of interest rather than beyond it - and use dynamic receive focusing. From Eq. 134, we note that this would mean a smaller F-number for a fixed aperture size - thus the spatial resolution would be better than our unfocused imaging technique. Some systems allow multiple transmit focal zones but at the cost of reduced frame rate. The technique of unfocused imaging allows one to reconstruct high quality images with improved echo SNR and spatial resolution compared to dynamic receive focusing in regions away from a fixed transmit focal zone - without a reduction in acquisition frame rate. We furthermore hypothesize that the technique could be implemented in realtime.

Although we did not implement shift-varying filtering techniques as suggested in the discussion on synthetic aperture interpretation, we anticipate that the shift-variant and 3-D nature of the beam could contribute to undesirable artifacts. The psf shape changes with depth, and also laterally near the edge of the transducer. To perform beamforming near the transducer edge, sub-aperture truncation must be performed. The resulting loss of aperture contributes to a laterally shift-variant psf - where longer and more energy rich tails are expected in the psfs. This is a phe-

nomenon that we have not accounted for - we have simply assumed approximate linear shift invariance. The shift-variance could be one reason why the unfocused filtered image quality is slightly worse than focal region images. Even without accounting for such details the matched filtering approach with one shift-variant filter has produced reasonable results. The success may be similar to filtering RF echo signals due to chirp excitations with mismatched filters[87] rather than matched filters. Future research should nevertheless aim to implement shift-varying filtering approaches that produce filtered image segments over smaller isoplanatic patches.

It is a challenge to the experimental design community to implement techniques of unfocused imaging that can indeed give comparable or superior images compared to other focusing-based techniques. The main requirements for accomplishing this goal are to find practical methods of implementing beamforming and apodization both on transmission and reception that minimize edge wave energy. Then, to find a fast method to implement the reconstruction processing in a shift-variant way.

8.5.1. Potential artifacts

Motion within the scan duration could produce significant artifacts using unfocused imaging techniques. Other groups, however, have successfully implemented related synthetic aperture techniques in vivo at very high frame rates. If scan times are fast enough, motion artifacts may be minimal.

Aberrations could be another source of artifacts. Aberrations due to differences in the speed of sound in tissue or due to refractive tissue interfaces could shift energy

away from the expected wavefronts. This, however, is also a problem for conventional focused imaging and perhaps more so for synthetic aperture techniques. We hypothesize that there is a tradeoff between curved wavefront length (SBP) and abberation strength. Understanding this better and applying it to in vivo imaging could be a topic of future work. Some novel combination of dynamic receive focusing and curved wavefront reconstruction methods may prove useful to this end and should be investigated further.

8.6. Conclusions

We have presented an imaging technique wherein wavefront curvature acts as a spatial code that can be decoded using matched filter or Wiener filtering techniques. The technique was motivated by fundamental theoretical work on the ideal observer for ultrasonic detection tasks. This theory provides a broader context for analyzing the effectiveness of systems designs for diagnostic tasks. In our investigation we found that the unfocused regions of focused beams contain equal or better lateral spatial bandwidth compared to the focal region. The reconstruction technique of spatial matched filtering is closely connected with synthetic aperture delay and sum techniques. Our approach is aimed at reconstructing RF data from focused transducers rather than point-sources. More transmit energy can thus be recovered with larger apertures. Edge waves in point-spread functions were seen to be a source of contrast degradation in both simulations and experiments. Apodization can potentially reduce the spatial sidelobe artifacts to acceptable levels. Importantly, image quality of

nearfield-reconstructed phantoms were better in terms of echo SNR and spatial resolution than dynamic receive focusing. The technique may be possible to implement in realtime with a linear array ultrasound system. It may also be possible to develop mechanically scanned single transducer or low channel count systems using this technique to offer much improved image quality in the unfocused regions of the images. In particular, our technique may have particular application to high-frequency ultrasonic imaging, where single element transducers are common and where arrays are expensive and hard to fabricate.

8.7. Appendix A

The spatio-temporal impulse response can be written as a series of temporal convolutions, (119). We begin our computation by using the Fresnel approximation to compute h_a using 126. As discussed, we assume that the Gaussian apodization is not severely truncated, and that $F - \sqrt{F^2 + r_2^2} \cong r_2^2/(2F)$. With these approximations, the integral in Eq. (126) can be written as

$$\begin{aligned} &\int dS\xi(\mathbf{r}) e^{i\frac{k}{2x_1}|\mathbf{r}|^2} e^{-i\frac{k}{x_1}\mathbf{x}\cdot\mathbf{r}} \\ &= \int e^{-[a_2 r_2^2 + b_2 r_2]} dr_2 \int e^{-[a_3 r_3^2]} dr_3 \end{aligned} \tag{148}$$

Where

$$\begin{aligned} a_2 &= \frac{1}{2\sigma_2^2} + j\frac{k}{2}\left(\frac{1}{x_1} - \frac{1}{F}\right) \\ b_2 &= -j\frac{k}{x_1}x_2 \end{aligned} \tag{149}$$

and

$$a_3 = \frac{1}{2\sigma_3^2} + j\frac{k}{2}\left(\frac{1}{x_1} - \frac{1}{F_e}\right) \tag{150}$$

Here we assume that $x_3 \approx 0$ so that the scatterers are approximately in the imaging plane. Additionally we can simplify the expression by assuming that the elevation focus F_3 is matched approximately to the imaging depth x_1 for simplicity. With this approximation, $a_3 = \frac{1}{2\sigma_3^2}$.

The integrals can be computed by completing the square. Note that

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) \tag{151}$$

hence the integrals we must solve have the form:

$$\begin{aligned}\int e^{-(ax^2+bx+c)}dx &= e^{-\left(c-\frac{b^2}{4a}\right)}\int e^{-a\left(x+\frac{b}{2a}\right)^2}dx \\ &= e^{-\left(c-\frac{b^2}{4a}\right)}\int e^{-ay^2}dy \\ &= e^{-\left(c-\frac{b^2}{4a}\right)}\sqrt{\frac{\pi}{a}}\end{aligned} \tag{152}$$

The Fresnel expansion of h_a thus becomes:

$$h_a(\mathbf{x}, f) \cong \frac{k}{j2\pi x_1}e^{jkx_1}\sqrt{\frac{\pi}{a_3}}\sqrt{\frac{\pi}{a_2}}e^{i\frac{k}{2x_1}x_2^2 - \frac{b_2^2}{4a_2}} \tag{153}$$

The right-most exponential term can be written as

$$e^{i\frac{k}{2x_1}x_2^2 - \frac{b_2^2}{4a_2}} = e^{-\Psi x_2^2} \tag{154}$$

with Ψ being a complex quantity $\Psi = \Psi_r + i\Psi_i$, where the real and imaginary parts are given by (129) and (130). As discussed, the real part defines a lateral Gaussian

envelope $e^{-\Psi_r x_2^2}$ for h_a, and the imaginary part defines a linear spatial frequency phase modulation $e^{-i(\Psi_i x_2)x_2}$, i.e. a baseband chirp.

We are interested in

$$h_{impulse}(\mathbf{x}, t) = \frac{\partial h_a(\mathbf{x}, t)}{\partial t} * h_a(\mathbf{x}, t). \tag{155}$$

To compute this, consider the frequency domain expression

$$h_{impulse}(\mathbf{x}, f) = jkc \times h_a^2(\mathbf{x}, f) \tag{156}$$

We then have

$$h_{impulse}(\mathbf{x}, f) \cong -jkc \times \left(\frac{k}{2\pi x_1}\right)^2 e^{j2kx_1} \frac{\pi}{a_3} \frac{\pi}{a_2} e^{-2\Psi x_2^2} \tag{157}$$

Before taking the inverse temporal Fourier transform of this, note that

$$\frac{1}{a_3} = 2\sigma_3^2 \tag{158}$$

and

$$\frac{1}{a_2} = \frac{\frac{1}{2\sigma_2^2} - j\frac{k}{2}\left(\frac{1}{x_1} - \frac{1}{F}\right)}{\left(\frac{1}{2\sigma_2^2}\right)^2 - \left(\frac{k}{2}\right)^2 \left(\frac{1}{x_1} - \frac{1}{F}\right)^2} \tag{159}$$

The real part of this is a Laplacian in k, thus has an inverse temporal Fourier transform of the form $e^{-\alpha|\tau|}$. The imaginary part also looks like a Laplacian but has an additional factor of jk in the numerator corresponding to a time-derivative in the temporal domain.

When the rightmost term in the denominator of 159 dominates, the approximation of neglecting $\left(\frac{1}{2\sigma_2^2}\right)^2$ is useful because the k^2 in the denominator cancels with a k^2 in

the numerator of 156 - simplifying the analysis.

$$h_{impulse}(\mathbf{x}, f) \cong -jkc \times \left(\frac{1}{x_1}\right)^2 e^{j2kx_1}\sigma_3^2 \times \tag{160}$$

$$e^{-2\Psi x_2^2}\left[\frac{1}{\sigma_2^2} - jk\left(\frac{1}{x_1} - \frac{1}{F}\right)\right] \times \frac{1}{\left(\frac{1}{x_1} - \frac{1}{F}\right)^2} \tag{161}$$

This can be written as:

$$\begin{aligned} h_{impulse}(\mathbf{x}, \omega) &\cong -\sigma_3^2\left(\frac{1}{x_1}\right)^2 \frac{1}{\left(\frac{1}{x_1} - \frac{1}{F}\right)^2} e^{-\frac{1}{\sigma_2^2}\frac{F^2}{(F-x_1^2)^2}x_2^2} \\ &\times \quad j\omega e^{j\omega\tau_F}\left[\frac{1}{\sigma_2^2} - j\frac{\omega}{c}\left(\frac{1}{x_1} - \frac{1}{F}\right)\right] \end{aligned} \tag{162}$$

where

$$\tau_F = \frac{2x_1}{c} - \frac{1}{x_1 - F}\frac{x_2^2}{c}. \tag{163}$$

Now proceeding with the inverse temporal Fourier transform, we have

$$\begin{aligned} h_{impulse}(t) &\cong -\sigma_3^2\left(\frac{1}{x_1}\right)^2 \frac{1}{\left(\frac{1}{x_1} - \frac{1}{F}\right)^2} e^{-\frac{1}{\sigma_2^2}\frac{F^2}{(F-x_1^2)^2}x_2^2} \\ &\times \quad \left[\frac{1}{\sigma_2^2} - \frac{1}{c}\left(\frac{1}{x_1} - \frac{1}{F}\right)\frac{d}{dt}\right]\frac{d}{dt}\delta(t - \tau_F) \end{aligned} \tag{164}$$

Although we could evaluate this using delta-function identities such as:

$$x^n\delta^{(n)}(x) = (-1)^n n!\delta(x) \tag{165}$$

we prefer to apply the temporal derivatives to the excitation of electro-mechanical coupling responses $h_{pulse}(t) \equiv h_y(t) * h_y(t) * v(t)$. In this way, the time-delay for the system impulse response is given completely and simply as τ_F:

$$\begin{aligned} h(\mathbf{x}, t) &\cong -\sigma_3^2\left(\frac{1}{x_1}\right)^2 \frac{1}{\left(\frac{1}{x_1} - \frac{1}{F}\right)^2} e^{-\frac{1}{\sigma_2^2}\frac{F^2}{(F-x_1^2)^2}x_2^2} \end{aligned}$$

$$\times \quad \left[\frac{1}{\sigma_2^2} - \frac{1}{c}\left(\frac{1}{x_1} - \frac{1}{F}\right)\frac{d}{dt}\right]\frac{d}{dt}h_{pulse}(t)$$

$$* \quad \delta(t - \tau_F) \tag{166}$$

Thus h can also be written as:

$$h(\mathbf{x}, t) = p(\mathbf{x}, t) * \delta(t - \tau_F(x)) \tag{167}$$

where $*$ is a temporal convolution and

$$\begin{aligned} p(\mathbf{x}, t) &= -\sigma_3^2\left(\frac{1}{x_1}\right)^2 \frac{1}{\left(\frac{1}{x_1} - \frac{1}{F}\right)^2} e^{-\frac{1}{\sigma_2^2}\frac{F^2}{(F-x_1^2)^2}x_2^2} \\ &\times \left[\frac{1}{\sigma_2^2} - \frac{1}{c}\left(\frac{1}{x_1} - \frac{1}{F}\right)\frac{d}{dt}\right]\frac{d}{dt}h_{pulse}(t). \end{aligned} \tag{168}$$

This time-delay factor can help us reduce spatial matched filtering operation for image reconstruction to a delay and sum procedure. Note that the second term of τ_F looks like a 2nd order Taylor expansion of $2[(x_1 - F) - \sqrt{(x_1 - F)^2 + x_2^2}]/c$ in x_2 about 0 (i.e. the Taylor series expansion that we used before with F replaced by $(x_1 - F)$).

9. EXPERIMENTAL TECHNIQUES FOR MEASURING IDEAL OBSERVER PERFORMANCE

In this chapter, we focus on measuring ideal observer performance. This is important for a number of reasons:

First, we need quantitative experimental performance validation of our ideal observer theoretical framework. It is possible to measure the 2-D psf of the system at different field locations, and to measure the Background and Noise Power spectrum to estimate the $GNEQ$ and thus the system performance. However, this assumes that the model we developed is accurate. We would like an experimental method that is blind to as many of the model assumptions as possible.

Second, we would like to be able validate specific design predictions, for example, to compare two system designs experimentally, with statistical significance. For example, we have predicted that the nearfield of a model system design is better than the focal region. We would like to be able to compare the detection performance in both of these regions. To do so with statistical significance is not trivial. A complicated problem - is how to put error bars on the measurements, and understand the uncertainties well enough to impact an experimental design for comparing performances. We would like to be able to know how to design phantoms (e.g. lesion size, contrast etc.) and to know how many images are needed to get error bars below a certain level to attain statistical significance in measured performance.

One of the challenges is that to make the error bars on detection performance very

small, one needs a large dataset of independent images where the lesion is at exactly the same location in all images. This is very difficult to realize experimentally with phantoms.

9.1. Techniques for Assessing Ideal Observer Performance

Here we review existing methods for measuring the ideal observer performance. We discuss their limitations for experimental techniques using phantom measurements.

9.1.1. Yes-No Experiments

One way of assessing the ideal observer performance is to have a large ensemble of signal present and signal absent images which are statistically independent speckle realizations. It is important that one have precise knowledge of the location of the lesion. For each image one would compute a test statistic value by (1) Wiener filtering (2) Integrating the resulting image-squared value over the lesion's spatial template. One may then histogram the resulting test statistic values for the signal present and signal absent cases. Given enough data, ROC analysis can be performed. Alternatively, the SNR of the test statistic may be estimated.

9.1.2. Two Alternative Forced Choice Experiments

Another technique for estimating the ideal observer performance is to use 2-Alternative Forced Choice experiments. In these experiments one needs a large ensemble of image pairs - one with a lesion present and one with a lesion absent. For each image pair the observer (in this case a computation) is forced to choose which image has the

lesion. In the case of the ideal observer, the test statistic is computed for each image in the pair, and the image with the greater test statistic value is scored as the image with the lesion present. The percent of correct decisions over a large number of image pairs is then a measure of performance that can be rigorously linked to the ideal observer SNR and the area under the ROC curve. One advantage of this technique is that the ideal observer test statistic does not need to be normally distributed for the SNR_{2AFC} to rigorously parameterize the ROC curve, and the ROC curve is always guaranteed to by symmetric about its diagonal midpoint.

2AFC techniques have been performed for ultrasonic detection tasks by Insana,[34] Abbey,[35] and Zemp.[84]

9.1.3. Experimental Challenges

Yes-No and 2AFC experiments are most appropriate when it is easy to simulate a large number of images with independent speckle realizations. We are interested in seing whether the estimated ideal observer performance from experimental data is as predicted from the 2-D speckle model. Experimental echo signals are due to a 3-D beam rather than a 2-D one. Also, the psfs in a real beam vary their shape with depth and even lateral position - factors which we have not accounted for in the model. Additionally, the underlying *pdf* of the object function may in fact be more complicated than a normal random process. If we had a way of measuring the ideal observer performance for real data (e.g. from a phantom) we could compare with the predicted values. Our objective is to validate the theoretical development before

using it for design purposes.

One challenge in using experimental data is obtaining a large number of statistically independent images that have the lesion location precisely known. Low contrast phantoms have been used to generate up to 100 independent images [ref], however, these experiments are not without challenges. Even so, 100 images is often not enough for demonstrating statistically significance for small performance differences. Often obtaining a large number of such images requires movement of the transducer or the inclusions, giving rise to location uncertainty, which has shown to produce quadratic rather than linear shapes of SNR vs contrast curves at low contrast levels [ref].

In this chapter we propose a method for estimating the ideal observer performance using much fewer images and where the lesion location is in an exactly know location - thus the technique is much more amenable to experimental measurement than past methods.

9.2. Error Analysis for Ideal Observer SNR Estimates using Y/N Experiments

In any technique that we use to estimate ideal observer performance, it will be important to understand the uncertainties, and to be able to predict the error bounds on estimates as this may impact experimental design. Here we consider the error bounds on estimates of means and variances of the log-likelihood test statistics. This will allow us to compute the relative uncertainty in SNR_I.

9.2.1. Uncertainties in the means and variances of the test statistic

We note that the uncertainty in the estimate of the mean of the test statistic is

$$\sigma_{\overline{\lambda}}^2 = \widehat{\sigma_{\lambda}^2}/N \tag{169}$$

where N is the number of independent images, and $\widehat{\sigma_{\lambda}^2}$ is the sample variance.

Before computing the error bounds on the variance, note that if $\lambda_1, \lambda_2, ..., \lambda_N$ are random test statistic values sampled from a normal distribution with parameters μ and σ^2 then the random variable

$$\widehat{var(\lambda|\pm)}/\sigma_{\lambda}^2 = \frac{\sum(\lambda_i - \overline{\lambda})_{\pm}^2}{\sigma_{\lambda}^2} \tag{170}$$

has χ^2 probability distribution with $\nu = N - 1$ degrees of freedom, given as:

$$\chi_{\nu}^2(x) \equiv \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{171}$$

Let $\chi_{\alpha,\nu}^2$, called a chi-squared critical value, denote the number on the measurement axis such that α of the area under the chi-squared curve lies to the right of $\chi_{\alpha,\nu}^2$. (These values can be computed using the inverse of the cumulative χ^2 distribution). Then a $100 \times (1 - \alpha)$ percent confidence interval for the variance σ^2 of a normal population has lower limit $(N-1)\widehat{\sigma_{\lambda}^2}/\chi_{\alpha/2,N-1}^2$ and upper limit $(N-1)\widehat{\sigma_{\lambda}^2}/\chi_{1-\alpha/2,N-1}^2$.

The error bounds on the means and variances of the test statistic will next be used to evaluate the relative uncertainty in SNR_I.

9.2.2. Uncertainty in SNR_I

In a previous chapter we saw that for low contrasts, $\sigma_\lambda^2 \equiv var(\lambda|+) \approx var(\lambda|-)$. Thus, the common variance in the denominator of SNR_I is given as

$$\sqrt{0.5(var(\lambda|+) + var(\lambda|-))} \approx \sigma_\lambda \tag{172}$$

and is estimated as the sample standard deviation $\widehat{\sigma}$.

Assuming that any correlation between sample means and variances is weak, (this will definitely be true if we use different sets of image realizations when we estimate the means than when we estimate the variances) we may write the uncertainty in SNR_I as

$$\frac{\delta(SNR_I)}{SNR_I} \approx \frac{\delta(\overline{\lambda}_+ - \overline{\lambda}_-)}{\overline{\lambda}_+ - \overline{\lambda}_-} + \frac{\delta(\widehat{\sigma_\lambda})}{\widehat{\sigma_\lambda}} \tag{173}$$

(Note also that a tighter bound could be formed by addition of variances in quadrature. We can denote this by replacing the + by a $\oplus$). Thus:

$$\frac{\delta(SNR_I)}{SNR_I} \approx \frac{2\widehat{\sigma_\lambda}/\sqrt{N}}{\overline{\lambda}_+ - \overline{\lambda}_-} + \epsilon_{\alpha,N} \tag{174}$$

where

$$\begin{aligned} \epsilon_{\alpha,N} &= max\left(\sqrt{\frac{N-1}{\chi^2_{1-\alpha/2,N-1}}} - 1, 1 - \sqrt{\frac{N-1}{\chi^2_{\alpha/2,N-1}}}\right) \\ &= \sqrt{(N-1)/\chi^2_{1-\alpha/2,N-1}} - 1 \end{aligned}$$

which is less for smaller sample variances but bounded by the rightmost term of 174. Curiously this can also be written as

$$\frac{\delta(SNR_I)}{SNR_I} \approx \frac{2}{\widehat{SNR_I}\sqrt{N}} + \epsilon_{\alpha,N} \tag{175}$$

The relative uncertainty in SNR_I is thus smaller for larger values of $\widehat{SNR}_I$ and for larger numbers of independent images N. Thus, in particular, large lesions will have smaller relative uncertainty than small lesions for a given number of images N.

Note that the experimental strategy for Y/N experiments could thus be quite different than those for 2AFC experiments, where error bounds are smallest for $SNR \approx 3$ or the percent correct $PC \approx 80\%$.

Interstingly, $\epsilon_{\alpha,N}$ does not depend on the speckle properties of the ultrasound scan, but rather tells only about the number of independent images. The properties of the speckle are inherent in the first term.

Of great practical interest is that the relative uncertainty in variance is independent of the size of lesion and independent of the speckle properties. This means that to estimate the relative uncertainty in variance, we may use many small patches to do so.

A plot of 90 percent confidence curves for the relative uncertainty in variance (i.e. $\epsilon_{\alpha,N}$ for $\alpha = 0.1$ and 0.25 and for values of $N = 1...300$ are shown in Fig. 47. Note that for $SNR_I > 3$, the $\epsilon_{\alpha,N}$ term dominates Eq. 174 for all values of N.

9.2.3. Target-Independent Performance

A valuable insight obtained from the analytic expression Eq. 115 is that the ideal observer SNR_I is separable into a product of target-dependent and system dependent quantities. A target independent figure of merit for low contrast detection tasks is thus given as:

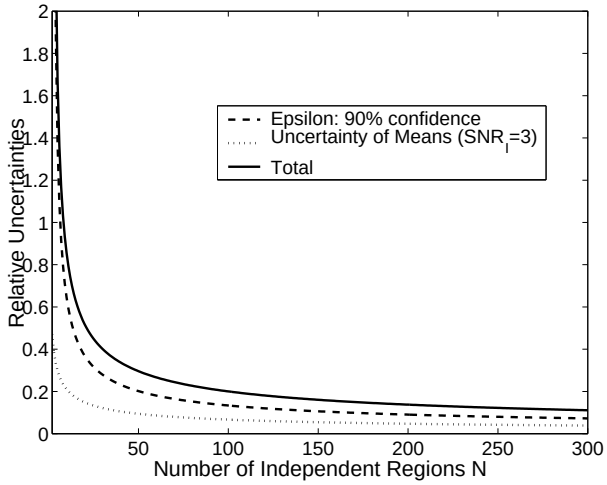

Figure 47. Relative Uncertainty of the Variance Estimate of the Test Statistic

$$\Upsilon \equiv \frac{SNR_I^2}{AC^2} = \int d\mathbf{u}\left(\frac{MTF(\mathbf{u})^2}{MTF(\mathbf{u})^2 + 1/SNR_e}\right)^2 \tag{176}$$

This suggests that we can choose any size of lesion (larger than a few speckle spots in radius) to measure performance, since the lesion size (and contrast - as long as it is low) do not affect the system performance metric Υ. Linearity of SNR_I^2 with lesion area has been tested using Monte Carlo simulation techniques.

The relative uncertainty in Υ is identical to the uncertainty in SNR_I:

$$\frac{\delta(\Upsilon)}{\Upsilon} \approx \frac{2}{\widehat{SNR}_I\sqrt{N}} + \epsilon_{\alpha,N} \tag{177}$$

This means that we may as well use experimental situations where SNR_I is large to estimate Υ. We propose using large lesion areas (but not necessarily large contrasts) to accomplish this. Again the experimental strategy is different than for 2AFC approaches.

9.2.4. Phantoms to Estimate the Mean Values of the Test Statistics

The differences in mean test statistic $\lambda_+ - \lambda_-$ between a lesion and its background can be estimated from a small number (e.g. 30) independent images - with minimum relative uncertainty for large lesions. (This can be accomplished without moving the transducer by making a phantom with a channel and have a insertable rod of slightly differing contrast than the background. The rod can then be rotated and translated to give many independent images). Such a phantom has been constructed and used for a number of different purposes.

9.2.5. Estimation of Variance using Homogenous Phantom Backgrounds

Approximating the variance inside and outside the lesion as nearly equal for low contrasts:

$$\widehat{\sigma}^2_{\lambda+} \approx \widehat{\sigma}^2_{\lambda-} \tag{178}$$

we may realize that we do not need to worry about obtaining many many low contrast lesion images in exactly the same location (this is part of the challenge of past experiments) to estimate the variance of the test statistic. Instead we can use homogeneous phantoms or phantom slurry composed of a finely blended suspension of phantom material. By stirring the phantom it is possible to obtain many independent image realizations. Additionally, it may be possible to use more than one data patch per image. By averaging this over M spatial patches per image (the patches do not need to be inside the lesion since low contrast variance in and outside the lesion are approximately equal), we can effectively increase the number of degrees of freedom to

$MN-1$ and hence significantly reduce the relative uncertainty in Υ. The uncertainty is thus given as:

$$\frac{\delta(\Upsilon)}{\Upsilon} \approx \frac{2}{\widehat{SNR}_I\sqrt{N}} + \epsilon_{\alpha,NM} \tag{179}$$

9.3. Estimators for SNR_I

Here we describe some analytic insights regarding how to estimate SNR_I by making use of ergodicity. With an ergodic assumption, many patches can be used per image, effectively increasing the number of degrees of freedom, and consequently reducing the relative error.

Our strategy is to obtain expressions for SNR_I and the relative uncertainty in SNR_I by developing estimators for the means and variances of the test statistic and characterizing their variance and bias. To do so it will be usefull to write the test statistic as:

$$\lambda = C\int_X w^2(\mathbf{x})d\mathbf{x} \tag{180}$$

where X is the region of the lesion, C is the object contrast, and w is the continuous form of the Wiener filtered signal $\mathbf{w} = \mathbf{K}_o^{-1}\mathbf{H}^t\mathbf{g}$.

9.3.1. Mean

Here we show how to construct an estimator for the mean of the test statistic

$$\overline{\lambda}_{\pm} = \langle C\int_X w^2(\mathbf{x})d\mathbf{x}|\pm\rangle = C\int_X \langle w^2(\mathbf{x})|\pm\rangle d\mathbf{x} \tag{181}$$

thus

$$\widehat{\overline{\lambda}}_{\pm} = \frac{1}{N}\sum_{i=1}^{N}\int_X w^2(\mathbf{x})d\mathbf{x}. \tag{182}$$

This expression states that the mean test statistic value can be estimated by averaging the test statistic value over N independent images. Its variance and bias decrease with increasing N.

9.4. Variance

At this point we derive estimators for the variance of the test statistic. Additional insights about ideal observer performance are also elucidated.

Note that

$$var(\lambda|\pm) = \langle\lambda^2|\pm\rangle - \overline{\lambda}_{\pm}^2 \tag{183}$$

and

$$\begin{aligned}
\langle\lambda^2|\pm\rangle &= \langle C^2\int_X w^2(\mathbf{x})d\mathbf{x}\int_X w^2(\mathbf{x}')d\mathbf{x}'|\pm\rangle \\
&= \int_X\int_X\langle w^2(\mathbf{x})w^2(\mathbf{x}')|\pm\rangle d\mathbf{x}d\mathbf{x}' \\
&= \int_X\int_X\left[\langle w^2(\mathbf{x})\rangle\langle w^2(\mathbf{x}')\rangle + 2\langle w(\mathbf{x})w(\mathbf{x}')\rangle^2\right]d\mathbf{x}d\mathbf{x}'
\end{aligned}$$

where the last line was obtained using a 4th order moment formula for Gaussian random variables. Thus,

$$var(\lambda|\pm) = 2\int_X\int_X\langle w(\mathbf{x})w(\mathbf{x}')\rangle^2 d\mathbf{x}d\mathbf{x}' \tag{184}$$

If we now assume that the signal corresponding to the lesion interior is approxi-

mately wide-sense stationary (should be true for low contrasts)

$$\widehat{var(\lambda|\pm)} = 2\int_X \int_X \widehat{R}^2_{ww}(\mathbf{x}-\mathbf{x}')d\mathbf{x}d\mathbf{x}' \tag{185}$$

where $\widehat{R}_{ww}(x)$ is the autocorrelation of the random process w. This is fortunate because we can easily estimate the autocorrelation from data without any assumptions about the data model. We can write the above equation as

$$\widehat{var(\lambda|\pm)} = 2\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \xi(\mathbf{x})\xi(\mathbf{x}')\widehat{R}^2_{ww}(\mathbf{x}-\mathbf{x}')d\mathbf{x}d\mathbf{x}' \tag{186}$$

where $\xi(\mathbf{x})$ is an window function that is one inside the lesion and zero outside. If we now make the substitutions $\tau = \mathbf{x} - \mathbf{x}'$ and $\zeta = \mathbf{x} + \mathbf{x}'$ then we have that

$$\widehat{var(\lambda|\pm)} = 2\int_{-\infty}^{+\infty} \widehat{R}^2_{ww}(\tau)f(\tau)d\tau \tag{187}$$

where

$$f(\tau) = \int_{-\infty}^{+\infty} \xi(\frac{\zeta+\tau}{2})\xi(\frac{\zeta-\tau}{2}) \tag{188}$$

which is purely a lesion geometry factor and has nothing to do with the system properties.

Eq. 187 is the estimator that we propose for estimating $var(\lambda|\pm)$. Later we show that its variance and bias decrease as the variance and bias of $\widehat{R}_{ww}$ decrease.

If $R_{ww}(\tau)$ tapers off to zero quickly so that most of the energy is contained in in the interval $|\tau| < \tau_c$ for some value τ_c and is slowly varying so that $f(\tau) \approx constant$ over this interval then

$$\widehat{var(\lambda|\pm)} \approx 2A\int_{-\infty}^{+\infty} \widehat{R}^2_{ww}(\tau)d\tau \tag{189}$$

where A is the area of the lesion, corresponding to $f(0)$. If $R_{ww}(\tau)$ has long tails with significant energy, the above approximation is not necessarily good.

9.4.1. Rectangular Lesion

To see this, consider a rectangular lesion, with lesion area bounded by a window function given as

$$\xi(x) = rect(x/X)rect(y/Y) \tag{190}$$

where X and Y are the width of the rect functions. Then,

$$f(\tau) = I_x(\tau_x)I_y(\tau_y) \tag{191}$$

where

$$I_x(\tau_x) = \int_{-\infty}^{+\infty} rect(\frac{\frac{\zeta_x+\tau_x}{2}}{X})rect(\frac{\frac{\zeta_x-\tau_x}{2}}{X})d\zeta_x \tag{192}$$

and I_y is given by a similar formula. The bounding equations for the rect functions: $-X/2 < \frac{\zeta_x+\tau_x}{2} < X/2$ and $-X/2 < \frac{\zeta_x-\tau_x}{2} < X/2$ imply that the intersecting area is over the interval $-X + |\tau_x| < \zeta_x < X - |\tau_x|$. Thus,

$$I_x = 2(X - |\tau_x|). \tag{193}$$

for $|\tau_x| < X/2$ and zero otherwise. This is approximately constant over small $|\tau_x|$ when X is very large. When X is small, $f(\tau)$ may emphasize long tails of $\widehat{R}_{ww}$ in 187, which increase the variance of the test statistic and degrades lesion visibility.

9.4.2. Circular Lesion

Here we assume that the lesions are circular in shape and that their area is bounded by a circular mask given by

$$\xi(\mathbf{r}) = circ(r/R) \tag{194}$$

where $r = |\mathbf{r}|$ and where R is the radius of the circle. Then

$$f(\tau) = \int circ(\frac{|\zeta + \tau|}{2R}) circ(\frac{|\zeta - \tau|}{2R}) d\zeta \tag{195}$$

We have the bounding equations: $(\zeta_x - \tau_x)^2 + (\zeta_y - \tau_y)^2 \leq (2R)^2$ and $(\zeta_x + \tau_x)^2 + (\zeta_y + \tau_y)^2 \leq (2R)^2$. The integral $f(\tau)$ represents the intersection area of two circles of radius 2R separated by distance $2|\tau|$. This area is:

$$\begin{aligned} f(\tau) &= 2\int_{y=|\tau|}^{y=2R} \int_{x=-\sqrt{(2R)^2-y^2}}^{x=+\sqrt{(2R)^2-y^2}} dxdy \\ &= 2\int_{y=|\tau|}^{y=2R} 2\sqrt{(2R)^2 - y^2} dy \end{aligned}$$

when $|\tau| < 2R$ and zero otherwise. From integral tables,

$$g_a(y) \equiv \int \sqrt{a^2 - y^2} dy = \frac{y\sqrt{a^2 - y^2}}{2} + \frac{a^2}{2} sin^{-1}(y/a) \tag{196}$$

thus, when $|\tau| < 2R$, we have that

$$f(\tau) = 4g_{2R}(y)|_{|\tau|}^{2R}. \tag{197}$$

Finally,

$$f(\tau) = (2R)^2 \left[\pi - 2sin^{-1}\left(\frac{|\tau|}{2R}\right)\right] - 2\tau\sqrt{(2R)^2 - |\tau|^2} \tag{198}$$

if $|\tau| < 2R$ and zero otherwise. As expected, when $|\tau| = 0$, the area is $\pi(2R)^2$, and when $|\tau| = 2R$, the area is zero. Sample curves are shown in Fig. 48.

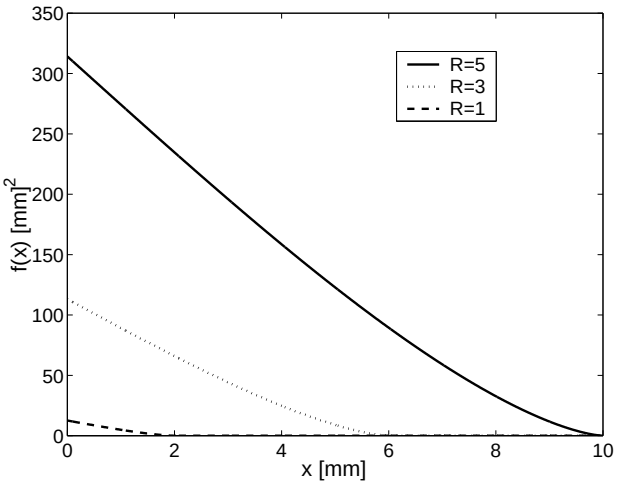

Figure 48. The function f in Eq. (198) shown for different values of lesion radius R.

Note that for small values of $|\tau|$, we have that

$$f(\tau) \approx A - \frac{4}{\sqrt{\pi}}\sqrt{A}|\tau| \tag{199}$$

where we have used a first order Taylor series approximation. Thus the variance is:

$$var(\lambda|\pm) = A\int R_{ww}^2(\tau)d\tau - \frac{4}{\sqrt{\pi}}\sqrt{A}\int R_{ww}^2(\tau)\tau d\tau \tag{200}$$

The second term is negligible when R_{ww}^2 falls to zero quickly, since it is greatest at zero, but then weighted by zero or small amounts. This reaffirms that for lesions containing many speckle spots, the variance of the test statistic is well approximated by the first term above. All of these hypotheses have yet to be validated and will need to be left for future work.

Because $var(\lambda|+) \approx var(\lambda|-)$ we can use homogeneous phantoms without inclusions to estimate the variance of the test statistic. The potential advantage of these estimators is that only small image patches can be used (big enough to sample most

of the energy in the autocorrelation $\widehat{R}_{ww}$) thus many more patches per image can be used - much larger than the lesion size.

We still need to derive expressions for the bias and variance of the estimators. We intend to show that the bias and variance of $\widehat{SNR}_I$ decreases with increasing numbers of data patches used to estimate $\widehat{R}_{ww}$.

9.4.3. Uncertainty of the variance estimate

We want to know the the relative uncertainty of the variance estimator $\widehat{\sigma}^2_{\lambda\pm}$ in Eq. (187), which we write as

$$\widehat{\sigma}^2_{\lambda\pm} = 2\frac{A}{(\sigma_w^2)^2}\int \widehat{\rho}^2(\mathbf{\Delta x}) f_o(\mathbf{\Delta x}) \tag{201}$$

where $\widehat{\rho}$ is the autocorrelation estimate normalized by its maximum value (i.e. $\widehat{R}_{ww}\sigma_w^2$), and $f_o = f/A$ where A is the maximum value of f.

We assume that can use N_p image patches to estimate this quantity. For each image patch n we construct an estimate $[\widehat{\sigma}^2_{\lambda\pm}]_n$, then average over all N_p patches to get the mean value of the estimate $\overline{\widehat{\sigma}^2_{\lambda\pm}}$. The standard deviation of this mean value is given as:

$$std\left(\overline{\widehat{\sigma}^2_{\lambda\pm}}\right) = \frac{std(\widehat{\sigma}^2_{\lambda\pm})}{\sqrt{N_p}} \tag{202}$$

Importantly, the relative uncertainty in the mean estimate of test statistic variance

$$\frac{std\left(\overline{\widehat{\sigma}^2_{\lambda\pm}}\right)}{\overline{\widehat{\sigma}^2_{\lambda\pm}}} = \frac{std(\widehat{\sigma}^2_{\lambda\pm})}{\sqrt{N_p}\overline{\widehat{\sigma}^2_{\lambda\pm}}} \tag{203}$$

is independent of the data variance σ_w^2, and can be viewed as being nearly independent of the lesion area. (Note the shape of f_o will still be dependent on the lesion area -

but the significance of the shape will be less for lesions several autocorrelation lengths in diameter). The relative uncertainty will depend in this situation on the ultrasonic speckle properties in a complex way.

The important point to be made is that one may use many data patches per image, thus N_p can be very large over a small number of independent images. Analytical expressions for the uncertainty in variance are difficult but discussed briefly below. For practical problems we recommend the more pragmatic approach of computing the standard deviation of N_p estimates of $\widehat{\sigma}^2_{\lambda\pm}$.

9.4.4. Analytical Expressions for the Uncertainty in Variance

In some cases it may be advantageous to predict the expected uncertainty in the variance of the test statistic if one knows analytically or computationally the expected autocorrelation profile of the RF data. For this reason we include a brief look into this complicated problem. Using 4th order Gaussian moment theorems, it can be shown that

$$var(\widehat{\sigma}^2_{\lambda\pm}) = 8\left(\frac{A}{(\sigma_w^2)^2}\right)\int \langle\widehat{\rho}(\Delta x_1)\widehat{\rho}(\Delta x_2)\rangle^2 f_o(\Delta x_1) f_o(\Delta x_2) d\Delta x_1 d\Delta x_2 \qquad (204)$$

Jenkins **??** gives a discussion regarding how to compute the covariance of the autocorrelation coefficients $\langle\widehat{\rho}[q,p]\widehat{\rho}[k,l]\rangle$. He shows that in one dimension

$$Cov[\widehat{\rho}(u_1),\widehat{\rho}(u_2)] = \frac{1}{(T-|u_1|)(T-|u_2|)}$$
$$\times \int_{-(T-u_1)}^{T-u_2} \phi(r)[\rho(r)\rho(r+u_2-u_1)\rho(r+u_2)\rho(r-u_1)]dr$$

where

$$\phi(r) = \begin{cases} T - u_2 - r & r \geq 0 \\ T - u_2 & -(u_2 - u_1) \leq r \leq 0 \\ T - u_1 + r & -(T - u_1) \leq r \leq -(u_2 - u_1) \end{cases} \tag{205}$$

and T is the correlation window length.

Future work should demonstrate how uncertainty in variance estimate can be reduced by using autocorrelation estimates that are averaged over many independent regions.

9.5. Bias of Estimator for Test Statistic Variance

The estimator of the test statistic variance is unfortunately biased. We would hope that the expectation of $\widehat{\sigma}^2_{\lambda\pm}$ would be equal to $\sigma^2_{\lambda\pm}$. There is however a bias that is hard to compute. In 1-D:

$$bias = \langle \widehat{\sigma}^2_{\lambda\pm} - \sigma^2_{\lambda\pm} \rangle = 2 \int \langle \widehat{R}^2_{ww}(\tau) - R^2_{ww}(\tau) \rangle f(\tau) d\tau \tag{206}$$

The bias of the variance of the autocorrelation sequence, thus depends on the expectation in the integral, which, when written as $\langle \widehat{R}^2_{ww}(\tau) \rangle - R^2_{ww}(\tau)$ can be identified as the variance of the autocorrelation estimates which can be written as:

$$\begin{aligned} var[\widehat{R}_{ww}(\tau)] &= \frac{1}{(T - |\tau|)^2} \int_{-(T-|\tau|)}^{(T-|\tau|)} (T - \tau - |r|) \\ &\times \left[\rho^2(r) + \rho(r + \tau)\rho(r - \tau) \right] dr \end{aligned}$$

We should note however, that if we first use many independent images to get a good estimate on the autocorrelation, then the variances (covariances) of the autocorrelation estimates become less and less, and consequently, the bias of the estimate of the variance of the test statistic is reduced. Specifically, averaging over M_p independent patches, the variance $var[\widehat{R}_{ww}(\tau)]$ is reduced by a factor of $(1/M_p)$ (or the standard deviation reduced by a factor $1/\sqrt{M_p}$). Consequently, the bias of the estimator for the test statistic variance can be reduced as low as desired with increasing numbers of independent data patches M_p.

10. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Here we outline the key contributions of this dissertation in the context of the existing literature. We also discuss limitations and offer suggestions for future work.

10.1. Rigorous Development of Linear Systems Model

Previous linear systems models of ultrasonic imaging have existed previously, but our work adds a rigorous connection to scattering theory, and emphasizes that the linear system should be written as

$$g(\mathbf{t}) = \int h(\mathbf{x}, \mathbf{t}) f(\mathbf{x}) d\mathbf{x} + n(\mathbf{t}) \tag{207}$$

rather than simply a temporal convolution as was previously assumed. One outcome of the analysis that was not well understood previous to our work is the difference between psfs and ssfs.

10.1.1. Difference Between psf and ssf

While for a 2-D imaging system, a psf is a 2-D image or 2-D RF echo signal from a point scatterer in a point in space, a ssf is the 3-D spatial field weighting that the system applies to the object function for a specific point in time. In much of our work we have assumed that a 2-D object and ssf is sufficient - an approximation that is reasonable as long as elevation beam effects are not significant. When a 2-D ssf is sufficient, there is a time-reversal symmetry that approximately relates the psf

and ssf over isoplanatic patches. An isoplanatic patch is a region on which linear shift invariance may be assumed locally. When the psf is larger than the size of the isoplanatic patch, the symmetry breaks down. ssfs, rather than psfs are required for evaluation of speckle statistics (e.g. spatial RF covariance matrices) which are in turn essential for image quality analysis and detection theory.

Through much of our thesis we have assumed the local linear shift-invariance and 2-D ssf approximation is sufficient. In particular, the local shift-invariance approximation enables Fourier decomposition techniques to be applied to the otherwise computationally foreboding problems of computing ideal observer test statistics and SNRs. An extension to the 3-D ssf has been shown to be computationally feasible, but local shift-invariance of each elevation slice of the psf must be assumed. It remains to be seen whether incorporating the 3rd dimension significantly impacts the prediction values of the ideal observer SNR. Such a study is recommended for future work.

10.1.2. Speckle Modeling in Pre- and Post- Focal Regions

A significant contribution of our linear system approach is the ability to model speckle statistics throughout the field - and not just at the focus as was done in previous work (e.g. Wagner). Existing notions that speckle spot size is indicative of spatial resolution exist in the ultrasonic community. While this is approximately true in the focal region, it is not true away from the focus. Our analysis shows that wavefront curvature can contribute to speckle narrowing. When wavefronts are flat as in the

focal region, coherence is maximized and speckle is broadest. A similar explanation was made previously by Trahey et. al. in analyzing speckle in the presence of aberrations. Speckle properties were found to be broadest when aberrations were least - and wavefront coherence was greatest.

The importance of modeling speckle away from the focal region was shown to be very important as it lead to some surprising discoveries about imaging in the pre- and post- focal regions, and the development of an imaging technique that can be viewed as spatial coding, and an extension of synthetic aperture techniques.

We have done some simple experimental validation of predicted speckle textures. In particular, we have approximated ssfs by time-reversed psfs, and done some token comparison experiments comparing speckle spot lengths, lesion contrast, and CNR's in simulated and real images, with modest results.

Future work could entail more precise experimental validation of the predictions of near- and farfield speckle autocorrelation functions. To do this we really need to use time-varying ssfs rather than linear shift-invariant psfs. Our theoretical work on speckle modelling showed the importance of ssfs for computing RF covariance matrices - and that using the time-reversed psf approximation can lead to small errors in computed autocorrelation curves and speckle correlation lengths. ssfs are hard to measure, since they require measuring a whole continuum of psfs. Thus the best option would be to have a realistic simulation of the system's ssfs. Current FIELD II simulations do not do beamforming like the real system: (1) Aperture

Modulation (2) Dynamic focusing using steered receive beams focused along A-scan lines. Complex and proprietary look-up tables currently implemented on the system would be needed in simulations to make exact the simulations fully realistic.

10.2. Development of Ideal Observer Model for Ultrasonic Detection Tasks

The main contribution of this dissertation is the derivation of an ideal observer model for ultrasonic detection tasks. Previous attempts at this (e.g. the Smith-Wagner model) did not fully account for RF correlations, shift variance, or electronic noise or were too cryptic for intuition and too difficult for computation (e.g. Clarkson and Barrett). Our approach fully incorporates the incomplete elements of previous models into our theoretical framework, while at the same time providing a simple analytical framework, that is also computationally trivial.

10.2.1. SKE Ideal Observer Test Statistics

An early contribution in our work was the derivation of ideal observer performance when the exact signal realization was known. The ideal observer SNR^2 for SKE tasks was then averaged over signal realizations. The result was an expression for task performance that decoupled the object properties such as contrast and lesion size and the system properties such as MTF and NPS. In particular, a generalization of noise-equivalent quanta was shown to be important for characterizing ultrasound system performance. The connection with NEQ and DQE is important in the context

of their widespread use in photon imaging modalities. Speckle spots replaced the role of photons as fundamental units of information.

10.2.2. SKS Ideal Observer Test Statistics

The log-likelihood ratio for SKS ultrasound detection tasks was derived as $\lambda = (\mathbf{K}_o^{-1}\mathbf{H}^t\mathbf{g})^t\,\mathbf{S}\,(\mathbf{K}_o^{-1}\mathbf{H}^t\mathbf{g})$. under a low contrast approximation. This suggests that the optimal RF processing strategy for low contrast detection is spatial matched filtering ($\mathbf{H}^t$) followed by pre-whitening $\mathbf{K}_o^{-1}$. The combined operation is equivalent to a spatial Wiener filter.

Note that spatial Wiener filtering is the optimal processing strategy for the ideal observer (whether this is the case for human observer is yet to be established) - independent of the system design. The system performance for detection tasks can be measured using SNR_I, or other figures of merit.

One of Abbey's main contributions to our effort was the power series method for signal present covariance inversion, which enables computation of the test statistic - a sufficient condition for Monte Carlo methods.

One of our key contributions is that we extended this idea to analytically compute the ideal observer SNR, and showed close agreement with the Monte Carlo methods. The analytical treatment allowed important intuitive insights, including connections with the Smith-Wagner theory. Additionally, important connections with Noise Equivalent Quanta and Detective Quantum Efficiency were again established, similar to the SKE theory. In the SKS theory, the $GNEQ^2$, rather than $GNEQ$

was integrated over the spatial frequency domain to give the relevant performance metric. Additionally, the SKS theory saw the proportionality of SNR_I^2 with AC^2, similar to the Smith-Wagner theory (whereas the SKE theory showed only linearity with contrast).

These developments are important for ultrasound system design. The analysis provides a toolbox that will allow us to push fundamental limits of ultrasound diagnostic performance.

Extension to a 3-D beam makes the $GNEQ$ analysis more complex, however, this may be unnecessary for problems where elevational properties of the beam are not expected to affect performance significantly. For example - in detecting lesions that are long in the elevation direction as is sometimes the case with cancers that grow along milk ducts.

10.3. Processing-Level Design and Detector Level Design

The reason we studies the ideal observer is to understand how to measure diagnostic information from the RF data. We call this level of design detector level design. Metrics such as the SNR of the ideal observer analysis gage system performance for tasks but the analysis does not however tell us *how* to design systems. We can however, from the analytical developments see intuitively what system features make a better design - and we can measure or predict the ideal observer performance for a given system design. This may allow us to compare several system designs and rank their performance for the diagnostic task needed.

The ideal observer test statistic suggested how to process the data in an optimal way for the ideal observer to maximize performance. This brings up the topic of image processing. Processing-level design generally refers to software development that acts to analyze the raw RF data for display to human observers. Processing algorithms cannot add diagnostic information to the data - but can make it more apparent to the human visual system. Performance of human observers viewing images created by a given set of algorithms can be measured using observer studies. Differences between human observer performance and the ideal observer suggest a role for improved image processing approaches.

Abbey showed that human observers are approximately 15% efficient relative to the ideal observer at performing low contrast detection tasks - when viewing unprocessed B-mode images. Previously human observers were thought to have efficiencies of closer to 50%. These estimates, however, were based on estimates of the Smith-Wagner theory, which ignored the feasibility of pre-whitening. The new estimate of efficiency is based on comparison with our newly proposed ideal observer model. The low efficiency suggested considerable room for improvement. Fortuitously, it was also shown that pre-envelope deconvolution greatly improved human observer performance - although there may be considerable room for additional improvement. Although the ideal observer test statistic suggested ways to process the raw data in a way optimal for the ideal observer, these algorithms may not be optimal for the human observer. This is why more research is needed to understand how to optimize image processing to the human visual system.

Sometimes the line between processing-level and detector-level design can be ambiguous and subjective. For example, for the technique of imaging with curved pre- or post- focal region wavefronts, we showed that the nearfield was even more informative than the focal region. In doing so we assumed that receive beamforming by delay and sum as well as apodization was applied to the signals from the individual channels. Beamforming can be viewed as a processing step rather than detector level design.

Future work should consider the raw data as signals from the individual channels. It is hypothesized that the ideal observer test statistic will suggest optimal processing strategies that can be applied to the channel data. Detector level design would thus come down to transmission beamforming design, and transmission and receive transducer design - and our framework is well poised to attack this problem in the near future.

10.3.1. Human Observers

Earlier we mentioned that differences between human observer performance and the ideal observer suggest a role for improved image processing approaches. Human observer studies can be expensive and time-consuming. For this reason, model observers that computationally simulate human observer performance could be invaluable. Additionally, a way of simulating random realizations of ultrasound images in a way that mimics real tissue and the real system are needed - indeed another challenge. So far our simulations have been shift-invariant and 2-D. In truth a shift-variant and 3-D model may be needed in some situations.

10.3.2. Task Specificity

Research in performance assessment is notorious for being very sensitive to changes in the task. This is true on many different levels. Even for the seemingly succinct task of lesion detection - there are many small variations in task definition - for example: the position and shape of the lesion may or may not be known - and if unknown - the position and shape may be characterized statistically. Ideal observers are sensitive to each of these cases and performance will differ. Although they are all detection tasks - the task detailed definitions can matter. We refer to this as task detail specificity.

Another concept to consider is that *completely different* tasks may have different system optimization curves. An example of completely different tasks could be low contrast lesion detection versus discrimination between an anechoic cyst and a high contrast hypo-echoic lesion. A system optimized for one task may not be - and probably will not be optimized for the other. Note that these differing tasks are still sometimes lumped together in the literature under the name detection. In clinical literature the word detection means: (1) saying whether or not there is an abnormality present (2) if the abnormality is a cyst or mass (3) if it is a mass one must decide whether it is benign (e.g. fibroadenoma) or malignant. Patients flagged as possessing lesions suspected of being maliganant are sent for biopsy. Mistakes in any one of the mentioned sub-tasks could be the cause of false positives of false negatives, and reduce the 'detection' performance. We envision a firm understanding of how to optimize each of these tasks - and have a programmable system that can optimize each stage

of the detection chain.

Back to the topic of task detail specificity - in our work we have assumed that the signal location was known exactly. Ideal observer detection performance is expected to drop considerably when location uncertainty is included in the framework - especially for low contrast detection tasks, since this is consistent with findings in other modalities. Similarly other sources of object and system uncertainty may additionally degrade ideal observer performance, since less and less information is known with more and more sources of uncertainty. A significant question is whether this matters for what one wants to do.

In particular, if one is interested in system design, a critical question is whether system rankings are preserved. For example, if one uses our framework to predict that one system design (System A) will be better than another (System B), then we factor in location and background uncertainty to formulate a new ideal observer prediction on system rankings ... is A still better than B? More research is needed to answer these questions for ultrasound as a modality. Factoring in additional sources of uncertainty into the ideal observer model is not necessarily trivial. Ideal analytically tractable likelihoods can be derived. In practice, however, marginalizing operations required to derive these probability densities are not analytically tractable. Computationally - the challenge is equally foreboding since integration over high dimensional spaces is often required. New computational tools are recently becoming available to tackle these problems by using Markov-Chain Monte Carlo methods. These methods avoid high-dimensional integrations by instead approximating marginalization integrations

by means of functions of Monte-Carlo generated data realizations. A pre-requisite of these techniques is to have a realistic probability model of object and system variability, as well as a way to computationally simulate realistic random realizations of image data in a fast way. In other modalities it was seen that inhomogeneous background variability degrades detection and influence system optimization [,]. Markov-Chain Monte Carlo techniques could be used to understand how inhomogeneous random background texture can influence system optimization. Our framework, nevertheless, is an important first step towards these more lofty goals.

10.4. Spatial Codes with Curved Wavefronts

Besides fundamental contributions, this dissertation includes an applied technique which we call spatial coding that offers potential improvements over both dynamic receive processing (the current method of choice on most systems) and traditional synthetic receive aperture techniques. Promisingly, the technique is likely computationally simple enough to implement in realtime.

10.5. Additional Suggestions for Future Work

More work needs to be done on the topic of experimental measurement techniques of the ideal observer performance of a system along with associated uncertainties. This is important for a number of reasons. First, we need quantitative experimental performance validation of our broad theoretical framework. Second, we would like to be able validate specific design predictions, for example, to compare two system

designs experimentally, with statistical significance. For example, we have predicted that the nearfield of a model system design is better than the focal region. We would like to be able to compare the detection performance in both of these regions. To do so with statistical significance is not trivial.

One problem is how to measure detection performance experimentally. Another - perhaps more complicated problem - is how to put error bars on the measurements, and understand the uncertainties well enough to impact an experimental design for comparing performances. We would like to be able to know how to design phantoms (e.g. lesion size, contrast etc.) and to know how many images are needed to get error bars below a certain level to attain statistical significance in measured performance.

One of the challenges is that to make the error bars on detection performance very small, one needs a large dataset of independent images where the lesion is at exactly the same location in all images. This is very difficult to realize experimentally with phantoms. For this reason we have proposed a technique that makes use of ergodicity - so that one or a few images can reliably model an entire ensemble of images. In these techniques, the measured autocorrelation is fundamental to estimating the variance of the test statistic of the log-likelihood. We have showed that the bias and variance of the estimate of test statistic variance decrease with increasing number of independent data patches used to estimate the spatial autocorrelation profile. We have also proposed that since the variance for the signal present and signal absent hypotheses should be approximately equal, it is sufficient to study homogeneous phantoms for the variance estimates. With an ergodicity assumption, it is thus possible to use one

or a few images to dramatically increase the number of degrees of freedom and hence shrink error bars. Yet another conclusion regarding minimizing the variance of the test statistic mean is that larger value of SNR_I are better. Consequently, it may be advantageous to use large lesions. Low contrasts are needed for the theory to still be valid. The measured SNR_I can be normalized by AC^2 to get the target-independent system performance.

It is possible to make a phantom with a channel for an insertable rod of differing contrast to be inserted, and translated or rotated to get 50 independent images of the lesion, where the lesion location is exactly known. The transducer does not need to be moved - hence the locational precision. Locational uncertainty in observer performance studies is known to give rise to erroneous effects [*]. The disadvantage of this method is that the background speckle realizations are unchanged from image to image. This may not matter, however, if the lesion is sufficiently large so that edge effects are negligible. Additionally, if the lesion is large enough, the uncertainty in the mean will be small - even for a small number of images - hence the severity of not having many background speckle realizations is reduced.

It may be anticipated that experimentally measured ideal observer performance may be less than theoretically predicted due to cumulative lack of model realism. The Wiener filter, for example, will be very sensitive to mistakes in the psf with the consequence of noise amplification and spatial sidelobe artifacts. Hence, again, there is even more reason to work towards more realistic simulations of system models.

10.6. Frontiers of Ultrasound System Design

Earlier it was noted that the ideal observer SNR only acts as a gage of design effectiveness - and does not necessarily tell one one how to design a system. Nevertheless, system designs can be evaluated for task performance - something that could not be done heretofore. An important conclusion to our work is to offer suggestions about how our developments can be applied to design better systems.

Future directions in ultrasonic imaging could include massively parallel receive acquisitions with wide transmit beams. Extensions of our curved wavefront imaging technique could be applied to this end. With faster and faster computing power evolving, one may be able to attain extremely high frame rates - and this could be applied to average significantly. This strategy intends to leverage on Moore's law which states that computing power will double approximately every 18 months. In using extensive averaging and in conjunction with axial and spatial codes one might be able to greatly improve the echo SNR - and hence go to higher and higher frequencies (thus improved spatial resolution) while preserving penetration. In the past, the amount of energy that could be used for diagnostic imaging was limited by standards. While standards still exist, these standards have recently changed to allow higher transmit energies - in part motivated by tissue harmonic imaging. Current guidelines require that acoustic output measurements merely be displayed on the system - and the discretion to go to higher acoustic intensities would be in the hands of the sonographer who is expected to manage risks responsibly. For high risk situations, increased acoustic output may be warranted. From this perspective,

this newfound latitude in standards could be used advantageously for certain classes of high risk disease.

One challenge in using ultrasound for a screening modality is that sonography is highly operator-dependent. Research could be done in scanning procedures and 3-D imaging technology to making the imaging less subjective. Another area of importance in mammography that could be extended to ultrasound is the area of computer aided diagnosis.

It may be that diagnostic B-mode imagine does not provide sufficient contrast compared with other competing modalities to distinguish between benign and malignant disease. Medical imaging frontiers lie in remote imaging of cellular pathology and biochemistry (sometimes referred to as molecular imaging), and other features that may distinguish malignant tissue from benign. Ultrasonic technique with promise include thermo- and photo-acoustic tomography which offer electromagnetic radio-frequency and optical contrast with the potential of ultrasonic spatial resolution. Additionally, ultrasonic strain imaging techniques aim to remotely map the mechanical properties of tissues. This offers considerable diagnostic potential since cancers are typically stiffer than the surrounding tissue. Much greater contrasts are expected compared with mammography. It is expected that improved algorithms and improved clinical instrumentation will move these research directions into clinically important techniques.

REFERENCES

1. *Surveillance, Epidemiology, and End Results (SEER) Cancer Statistics Review 19731999*, http://cis.nci.nih.gov.
2. M. F. Insana, L. T. Cook, M. Bilgen, P. Chaturvedi, and Y. Zhu, "Maximum-likelihood approach to strain imaging using ultrasound,,"
3.
4. Gara
5. N. M. Lindor and M. H. Greene, "The concise handbook of family cancer syndromes: Mayo familial cancer program.," *J Natl Cancer Inst* **90**, pp. 1039–71, 1998.
6. ?
7. M. S. Sabel and E. D. Staren, "Innovations in breast imaging: How ultrasound can enhance the early detection of breast cancer," *Medscape Womens Health* **2**, p. 1, 1997.
8. S. C. Chen, Y. C. Cheung, Y. F. Lo, M. F. Chen, T. L. Hwang, C. H. Su, and S. Hsueh, "Sonographic differentiation of invasive and intraductal carcinomas of the breast," *Br. J. Radiol.* **76**(909), pp. 600–4, 2003.
9. Irwig
10. T. M. Kolb, J. Lichy, and J. H. Newhouse, "Occult cancer in women with dense breasts: detection with screening us–diagnostic yield and tumor characteristics," *Radiology* **207**, pp. 191–199, 1998.
11. H. H. Barrett, "Objective assessment of image quality: Effects of quantum noise and object variability," *J. Opt. Soc. Amer. A* **7**, p. 12661278, 1990.
12. H. V. Trees, *Detection, Estimation, and Modulation Theory*, Wiley, New York, 1971.
13. H. H. Barrett, C. K. Abbey, and E. Clarkson, "Objective assessment of image quality. iii. roc metrics, ideal observers, and likelihood-generating functions," *J. Opt. Soc. Amer. A* **15**, p. 15201535, June 1998.
14. Green and Swets.
15. Gonzalez and Woods, ********Image Processing.*
16. M. Lukacs, M. Sayer, and S. Foster, "Single element and linear array pzt ultrasound biomicroscopy transducers," in *Proc. IEEE Ultrason. Symp.*, pp. 1709 – 1712, Oct. 1997.

17. R. F. Wagner, M. F. Insana, and D. G. Brown, "Statistical properties of radio-frequency and envelope-detected signals with applications to medical ultrasound," *J. Opt. Soc. Amer. A* **4**, p. 910922, 1987.

18. R. Y. Chiao and X. Hao, "Coded excitation for diagnostic ultrasound: A system developer's perspective," in *Proc. IEEE Ultrason. Symp.*, pp. 437–448, Oct. 2003.

19. P.-C. Li and J.-J. Huang, "Efficient dynamic focus control for three-dimensional imaging using two-dimensional arrays," *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.* **49**(9), pp. 1191 – 1202, Sept. 2002.

20. B. A. J. Angelsen.

21. ?

22. A. T. Fernandez, K. L. Gammelmark, J. J. Dahl, C. G. Keen, R. C. Gauss, and G. E. Trahey, "Synthetic elevation beamforming and image acquisition capabilities using an 8 x 128 1.75d array," *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.* **50**(1), pp. 40–57, Jan. 2003.

23. H. H. Barrett and K. J. Myers, *Foundations of Image Science*, John Wiley and Sons, 2003.

24. R. Shaw, "The equivalent quantum efficiency of the photographic process,,"

25. A. E. Burgess, "The rose model, revisited," *J. Opt. Soc. Am. A.* **16**, pp. 633–646, 1999.

26. Tanner and Birdsal.

27. R. F. Wagner and D. G. Brown, "Unified snr analysis of medical imaging systems," *Phys. Med. Biol.* **30**, pp. 489–518, 1985.

28. R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound b-scans, ieee trans. sonics ultrason.," **30**, p. 156163.

29. S. W. Smith, R. F. Wagner, J. M. Sandrik, and H. Lopez, "Low contrast detectability and contrast/detail analysis in medical ultrasound," **30**, pp. 164–173, May 1983.

30. H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Meyers, "Objective assessment of image quality. ii. fisher information, fourier crosstalk, and figures of merit for task performance," *J. Opt. Soc. Amer. A* **12**, p. 834852, May 1995.

31. Rolland and Barrett

32. K. J. Myers, J. P. Rolland, H. H. Barrett, and R. F. Wagner, "Aperture optimization for emission imaging: effect of a spatially varying background.," *J. Opt. Soc. Am. A.* **7**(7), pp. 1279–1293, 1990.

33. J. W. Goodman, *Introduction to Fourier Optics, 2nd ed.*, McGraw-Hill, New York, 1996.

34. M. F. Insana and T. J. Hall, "Visual detection efficiency in ultrasonic imaging: A framework for objective assessment of image quality," *J. Acoust. Soc. Am.* **95**(4), pp. 2081–2090, April 1994.

35. C. K. Abbey, R. J. Zemp, and M. F. Insana, "Pre-envelope deconvolution for increased lesion detection efficiency in ultrasonic imaging," in *Proc. SPIE*, vol. 5034, 2003.

36. E. Clarkson and H. H. Barrett, "Approximations to ideal-observer performance on signal-detection tasks,,"

37. G. E. Tupholme, "Generation of acoustic pulses by baffled plane pistons," *Mathematika* **16**, p. 209224, 1969.

38. P. R. Stepanishen, "The time-dependent force and radiation impedance on a piston in a rigid infinite planar baffle," *J. Acoust. Soc. Amer.* **49**, p. 841849, Mar. 1971.

39. P. R. Stepanishen, "Transient radiation from pistons in a infinite planar baffle," *J. Acoust. Soc. Amer.* **49**, p. 16271638, May 1971.

40. t. . P R Stepanishen

41. P. R. Stepanishen, "Pulsed transmit/receive response of ultrasonic piezoelectric transducers," *J. Acoust. Soc. Amer.* **69**, p. 18151827, June 1981.

42. J. C. Gore and S. Leeman, "Ultrasonic backscattering from human tissue: A realistic model," *Phys. Med. Biol.* **22**, p. 317326, 1997.

43. M. Fatemi and t. . A C Kak

44. J. A. Jensen, "A model for the propagation and scattering of ultrasound in tissue,,"

45. W. F. Walker and G. E. Trahey, "The application of k-space in pulse echo ultrasound," *IEEE Trans. Ultrason., Ferroelect. Freq. Contr.* **45**, p. 541558, May 1998.

46. W. F. Walker, "The significance of correlation in ultrasound signal processing," in *Proc. IEEE*, vol. 4325, p. 159171, Feb. 2001.

47. L. A. Chernov, *Wave Propagation in a Random Medium*, McGraw-Hill, New York.

48. L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics, 4th ed*, Wiley, New York, 2000.

49. M. F. Insana and D. G. Brown, "Acoustic scattering theory applied to soft biological tissues," in *Ultrasonic Scattering in Biological Tissues*,

50. P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, McGraw-Hill, New York, 1968.

51. D. G. B. M F Insana, R F Wagner and T. J. Hall, "Describing small-scale structure in random media using pulseecho ultrasound," **87**, p. 179192, Jan. 1990.

52. M. Ueda and H. Ichikawa, *Analysis of an echo signal reflected from a weakly scattering volume by a discrete model of the medium*, vol. 70, Dec. 1981.

53. J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers,,"

54. P. T. Christopher and K. J. Parker, "New approaches to the linear propagation of acoustic fields,,"

55. G. E. Tupholme, "Generation of an axisymmetrical acoustic pulse by a deformable sphere,,"

56. J. A. Jensen, *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach*, Cambridge Univ. Press, New York, 1996.

57. F. Kallel, M. Bertrand, and J. Meunier, "Speckle motion artifact under tissue rotation,,"

58. M. F. Insana and L. T. Cook, "Bioelasticity imaging: System design,,"

59. S. W. Flax and M. ODonnell, "Phase-abertation corrections using signals from point reflectors and diffuse scatterers: Basic principles," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.* **35**, p. 758767, Nov. 1988.

60. S. W. Smith, G. E. Trahey, S. M. Hubbard, and R. F. Wagner, "Properties of acoustical speckle in the presence of phase aberration. ii. correlation lengths (medical application)," *Ultrason. Imag.* **10**, p. 2951, Jan. 1988.

61. J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, Wiley-Interscience, New York, 1971.

62. M. ODonnell, "Coded excitation system for improving the penetration of real-time phased-array imaging systems," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.* **39**, p. 341351, May 1992.

63. M. E. Anderson, "Multi-dimensional velocity estimation with ultrasound using spatial quadrature," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.* **45**, p. 852861, May 1998.

64. J. A. Jensen and P. Munk, "A new method for estimation of velocity vectors,,"

65. D. Middleton, *An Introduction to Statistical Communication Theory*, Penninsula Publishing, Los Altos, CA, 1987.

66. G. H. Golub and C. F. VanLoan, *Matrix Computations, 2nd Ed.*, Johns Hopkins University Press, Baltimore, Maryland, 1989.

67. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, http://www.netlib.org/etemplates/, 2000.

68. H. H. Barrett, J. L. Denny, H. C. Gifford, and C. K. Abbey, "Generalized neq: Fourier analysis where you would least expect to find it," in *Proc. SPIE*, vol. 2708, pp. 41–52.

69. J. A. Jensen and P. Gori, "Spatial filters for focusing ultrasound images," *Proc. IEEE Ultrason. Symp.* , pp. 1507–1511, 2001.

70. R. J. Zemp, C. K. Abbey, and M. F. Insana, "Linear system models of ultrasound imaging: Application to signal statistics," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.* **50**, pp. 642–654, 2003.

71. R. M. Gray, *Toeplitz and Circulant Matrices: A Review,*.

72. A. T. Stavros, D. Thickman, C. Rapp, M. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions," *Radiology* **196**, pp. 123–124, 1995.

73. R. J. Zemp, C. K. Abbey, and M. F. Insana, "Generalized neq for assessment of ultrasound image quality," in *Proc. SPIE*, vol. 5030, pp. 391–402, 2003.

74. R. J. Zemp, C. K. Abbey, and M. F. Insana, "Fundamental performance metrics and optimal image processing strategies for ultrasound systems," in *Proc. IEEE Ultrason. Symp.*, 2002.

75. A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes, 4th Ed.*, McGraw-Hill.

76. B. Haider, P. Lewin, and K. E. Thomenius, "Pulse elongation and deconvolution filtering for medical ultrasonic imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.* **45**(1), Jan. 1998.

77. A. Burgess, "Image quality, the ideal observer, and human performance of radiologic decision tasks," *Acad. Radiol.* **2**(6), pp. 522–526, June 1995.

78. M. A. Kupinski, J. W. Hoppin, E. Clarkson, and H. H. Barrett, "Ideal observer computation using markov chain monte carlo,,"

79. A. S. for Real-time Display of Thermal and R. . Mechanical Acoustic Output Indices on Diagnostic Ultrasound Equipment 1998.

80. B. Ward, A. C. Baker, and V. F. Humphrey, "Nonlinear propagation applied to the improvement of resolution in diagnostic medical ultrasound," *J. Acoust. Soc. Am* **101**(1), pp. 143–154, Jan. 1997.

81. R. J. Zemp, J. Tavakkoli, and R. S. C. Cobbold, "Computationally efficient modeling of nonlinear ultrasound propagation," *J. Acoust. Soc. of Am.* **113**(1), pp. 139–152, Jan. 2003.

82. Add

83. J. A. Jensen, "Coded excitation,"

84. R. J. Zemp, M. D. Parry, C. K. Abbey, and M. F. Insana, "Ultrasound detection theory,"

85. Nikolov and Jensen, "Synthetic aperture imaging with virtual sources,"

86. G. Trahey..., "???,"

87. T. Misaridis, "???,"