Published in final edited form as:
IEEE Trans Med Imaging. 2005 March ; 24(3): 300–310.

Detection Performance Theory for Ultrasound Imaging Systems

Roger J. Zemp[Student Member, IEEE],
Department of Biomedical Engineering, University of California Davis, Davis, CA 95616 USA. He is now with the Optical Imaging Laboratory, Department of Biomedical Engineering, Texas A&M University, 233 Zachry Engineering Center, 3120 TAMU, College Station, TX 77843 USA (e-mail: rzemp@oilab.tamu.edu).

Mark D. Parry,
Department of Biomedical Engineering, University of California Davis, Davis, CA 95616 USA.

Craig K. Abbey[Member, IEEE], and
Department of Biomedical Engineering, University of California Davis, Davis, CA 95616 USA and also with the Department of Psychology, University of California Santa Barbara, Santa Barbara, CA 93101 USA.

Michael F. Insana[Member, IEEE]
Department of Biomedical Engineering, University of California Davis, Davis, CA 95616 USA. He is now with the Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: mfi@uiuc.edu).

Abstract

A rigorous statistical theory for characterizing the performance of medical ultrasound systems for lesion detection tasks is developed. A design strategy for optimizing ultrasound systems should be to adjust parameters for maximum information content, which is obtained by maximizing the ideal observer performance. Then, given the radio-frequency data, image and signal processing algorithms are designed to extract as much diagnostically relevant information as possible. In this paper, closed-form and low-contrast approximations of ideal observer performance are derived for signal known statistically detection tasks. The accuracy of the approximations are tested by comparing with Monte Carlo techniques. A metric borrowed and modified from photon imaging, Generalized Noise Equivalent Quanta, is shown to be a useful and measurable target-independent figure of merit when adapted for ultrasound systems. This theory provides the potential to optimize design tradeoffs for detection tasks. For example it may help us understand how to push the limits of specific features, such as spatial resolution, without significantly compromising overall detection performance.

Index Terms

Cancer; decision theory; image quality; speckle

I. INTRODUCTION AND MOTIVATION

Ultrasound systems are often characterized by parameters such as spatial resolution, contrast resolution, and echo signal-to-noise. Because there are inherent tradeoffs among these features, a judicious balance must be made that is application dependent. How to do this is

Correspondence to: Roger J. Zemp.

frequently unclear. This paper outlines a method to optimize system design tradeoffs in a task specific way. From this perspective, we define image quality specifically as the ability of the system to perform a given diagnostic task. In this paper, we focus on detection tasks. One application of this theory could be breast cancer detection. In a recent screening study [1], ultrasound in combination with mammography was shown to significantly improve the detection of breast masses compared to mammography alone, especially in women with dense breasts. Because ultrasound exams are safe, quick, and relatively inexpensive, new technology optimized for detection may enhance the role of ultrasound in screening programs as an adjunct to mammography or as a stand-alone screening modality.

Our methodology is to model essential features of the system and the body with a statistical framework. A linear systems model can accurately represent radio-frequency (RF) echo signals given object variability and electronic noise. Equipped with the statistical model, we then approach the problem from a statistical detection perspective.

Smith and Wagner [2]–[4] developed a statistical model for ultrasonic image quality that provided new insights into the detection process even though it ignored electronic noise and assumed object functions that generate fully developed speckle, shift-invariant focused imaging systems, and large low-contrast lesions. They also studied B-mode images and the squared envelope signal rather than the RF signal. We hope to relax some of these assumptions and include noise in the analysis. Our goal is to use the theory to develop systems that maximize the information content of the raw RF signal specifically for lesion detection and understand the role of phase information discarded by B-mode systems.

The analysis developed here provides simple but powerful design equations that can be used to optimize ultrasound systems for detection tasks in a fundamentally new way. It will allow us to explore the merits of new techniques for breast cancer detection and other applications.

II. BACKGROUND

A. Task-Based Approaches to System Design and Image Quality Assessment

Our approach to studying image quality for ultrasound systems is quantitative and specific to detection tasks. We define our detection task as follows: given an image or raw RF echo signal data, decide between two hypotheses H_1 and H_0. The signal present hypothesis H_1 is that there is a lesion present. The null hypothesis H_0 is that there is no lesion. This type of task is often called a YES/NO (YN) task [5]. A related task that we will sometimes consider is a 2-Alternative Forced Choice (2AFC) experiment. The task is to consider two image datasets: one with the lesion present (H_1) and one with the lesion absent (H_0) and decide which of the two images has the lesion. 2AFC tasks are rarely encountered clinically but they are useful for controlled observer performance experiments.

Task assessment can take place either at the detection or display stage of image formation [6]. *Detector-level* assessments evaluate designs based on the information content of the echo signals recorded. Factors influencing information content include transducer design, beamformation, aberration correction, coded pulse transmission, contrast agents, and prebeamformer and postbeamformer electronics. *Display-level* assessments seek to evaluate the human observer's accessibility to information. Processing methods may enhance display-level performance but only to the extent that the diagnostic information is originally in the raw echo signal (hence, the importance of detector level development). Processing methods include filtering, deconvolution, compounding, color-flow processing, elasticity estimation, etc. This paper primarily focuses on detector-level development. To do this we aim to optimize the performance of the ideal observer [7].

B. Observers for Detection Tasks

Observers include trained humans and numerical algorithms that make decisions based on an examination of the data. Mathematical observers use a scalar decision function of the data vector **g** as a *test statistic* $t(\mathbf{g})$. In this paper, we consider the data vector **g** to be the beamformed RF echo signal. B-mode systems process the echo signals to form A-scan lines that form the digital image **b** displayed by the system. A numerical observer makes a binary decision by comparing the test statistic with a scalar threshold t_c. If $t(\mathbf{g}) > t_c$, decision D_1 (signal present) is made, otherwise decision D_0 is made. For detection tasks, human observer studies are conducted to measure detection performance, often at great expense. Mathematical observers are an efficient substitute for humans if their performance can be used to predict human performance.

Performance is quantified with a receiver operating characteristic (ROC) curve [5]. When the test statistic is normally distributed, the ROC curve is parameterized by the detection signal-to-noise ratio (SNR)

$$\mathrm{SNR}_t = \frac{\bar{t}_1 - \bar{t}_0}{\sqrt{\frac{1}{2}(\sigma_1^2 + \sigma_0^2)}} \tag{1}$$

where the quantities in the numerator and denominator are the mean and variance of the test statistic conditioned on the signal present and signal absent hypotheses.

C. Ideal Bayesian Observer and Its Role in System Design

In this paper, we are concerned with the *ideal observer*, which acts not on the envelope detected B-mode image **b**, but rather on the raw RF echo signals **g**. The ideal observer is the optimum decision-function $t = \lambda(\mathbf{g})$ given the data. It is ideal in the Bayesian sense that λ yields the largest true-positive probability for any false-positive level. The test statistic of the ideal observer is the log-likelihood ratio [5], [8]

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

We would like to be able to model (and measure) the performance of the ideal observer decision function because maximizing the ideal observer performance means maximizing the diagnostic information content in the raw data [8] (*detector-level development*). It is the role of signal and image processing algorithms to make the information in the raw data maximally apparent to human observers (*display- level processing*).

III. THEORY

A. 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 (in-phase and quadrature) echo signal **g** can be modeled as a linear system of the object function [9]. IQ data ($\mathbf{g}_I$, $\mathbf{g}_Q$) is formed by mixing RF echo signals 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 by rectifying and low-pass filtering: $\mathbf{b} = \mathrm{LPF}(|\mathbf{g}|)$. In computing the log-

likelihood ratio, Smith and Wagner considered independent speckle spots of the B-mode or intensity signal rather than correlated pixels of the RF data. In this paper, the log-likelihood calculation is based on the pre-envelope (RF or IQ) signal to avoid losing information during B-mode processing.

A solution to the small-amplitude acoustic wave equation in scattering media has led to a linear systems model of ultrasonic imaging [9], represented as $\mathbf{g} = \mathcal{H} f + \mathbf{n}$, where $\mathbf{g}$ is a vector of digitized echo signals, $f(\mathbf{x})$ is the object function defined over spatial locations $\mathbf{x}$, $\mathcal{H}$ is a continuous-to-discrete integral operator that maps object functions into data space and $\mathbf{n}$ represents electronic and quantization noise. The linear model is capable of describing speckle texture throughout the beam profile—not just in the focal region. The model explains, for example, why nearfield speckle is finer than at the focus—even though point-spread functions are broader [9]. For a $N \times M$ matrix of recorded echo signals, $\mathbf{g}$ is a $NM \times 1$ RF column vector formed by vertically concatenating the RF image columns. For simplicity we consider two-dimensional (2-D) object and system models, although three-dimensional (3-D) models could be considered. For 3-D beams, the echo signal $\mathbf{g}$ can be considered a coherent sum of echo signals from a vertical ensemble of 2-D object functions f_i such that $\mathbf{g} = \sum_i \mathcal{H}_i f_i + \mathbf{n}$, where $\mathcal{H}_i$ is the 2-D slice through the spatial sensitivity function (ssf) at elevation i.

The sum can again be written as a linear system $\mathbf{g} = \mathcal{H} f + \mathbf{n}$ by redefining $\mathcal{H}$ and f. The 2-D model assumes that one elevational slice of the ssf is adequate. A 2-D model is presently justifiable—especially when considering targets with elevational depths greater than the elevational beamwidth.

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

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

The isoplanatic assumption leads to $NM \times N'M'$ block-Toeplitz $\mathbf{H}$ matrices which may be approximated as block-circulant. The error in this approximation is expected to be small if the psf is compact [10]. This matrix notation will allow us accessibility to a wide range of statistical tools, matrix identities, and fast Fourier computational techniques. This simple approach helps us focus on optimization of the axial and lateral properties of the beam. Similar to Abbey [11], we estimated system and noise parameters for the statistical model from a 7.2-MHz transducer using a Siemens Elegra Ultrasound system by scanning a graphite-agar phantom. The 2-D RF power spectrum was estimated from 10 spatially independent recorded echo fields near the transmit focal region. Dynamic receive and aperture growth1 were used so that the speckle is very uniform over the image (large isoplanatic patches). The pulse parameters were fit to a 2-D Gaussian function. The measured echo SNR of these images was 47 dB (defined as the peak value of the signal power spectral density divided by the noise power spectral density level),2 and the estimated axial and lateral pulse widths were specified by the Gaussian parameters $\sigma_x = 0.08$ and $\sigma_y = 0.15$ mm, respectively. The axial and lateral dimensions of a pixel are 0.054 mm by 0.12

[1]Dynamic receive processing allows one to focus at each receive location—even when only one transmit focus is possible. Aperture growth is a way of dynamically changing the receive aperture size to maintain an approximately constant receive F-number.
[2]Under the definition of signal power divided by noise power, the echo SNR is 43.2 dB.

mm, respectively. Fig. 1 shows a representative schematic of the linear system model of image formation. Fig. 2 compares images formed from simulated and measured echo data from a uniform Rayleigh scattering phantom.

B. Signal Known Statistically

By modeling essential features of the system and the tissue, a statistical detection theory can be formulated for ultrasonic tumor imaging. Although shape, echogenicity, spatial heterogeneity, margin characteristics, and posterior acoustic shadowing are important diagnostic features in applications such as breast cancer sonography [12] we focus on echogenicity as shown in Fig. 1. Furthermore, although the role of breast sonography so far has been as an adjunct to mammography and as a means of discriminating lesions as benign or malignant, we choose to focus on ultrasound as a detection modality. This paper may be considered as a stepping stone to more general discrimination tasks since the theoretical descriptions are similar. We presently consider a signal-known-statistically (SKS) detection task. Known features of the SKS task are 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 multivariate normal (MVN) 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

$$\begin{aligned} \mathrm{pr}[\mathbf{g}\,|\,H_1] &= \mathrm{MVN}(\mathbf{0}, \mathbf{K}_1) \\ \mathrm{pr}[\mathbf{g}\,|\,H_0] &= \mathrm{MVN}(\mathbf{0}, \mathbf{K}_0). \end{aligned} \tag{4}$$

In a manner similar to previous work [11], [13]–[15], the signal present and signal absent covariance matrices have the form $\mathbf{K_g}=\mathbf{H}\mathbf{K_f}\mathbf{H}^t+\sigma_n^2\mathbf{I}$, where $\mathbf{K_f}$ represents the covariance of the object and represents the additive noise. We represent a lesion as an object region where the variance differs from the surrounding region (see Fig. 1). In general, the signal present covariance matrix $\mathbf{K}_1 \equiv \mathbf{K}_{\mathbf{g}|H_1} = \langle \mathbf{g}\mathbf{g}^t|\boldsymbol{H_1}\rangle$ can be thought of as the sum of the background covariance and a differential signal covariance $\mathbf{K}_1 = \mathbf{K}_0 + \boldsymbol{\Delta}\mathbf{K}$, where $\boldsymbol{\Delta}\mathbf{K} = \mathbf{K}_1 - \mathbf{K}_0$ and where $\mathbf{K}_0 \equiv \mathbf{K}_{\mathbf{g}|H_0}$.

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

The log-likelihood ratio is the decision function of the ideal observer. From (2) and (4) it may be written in terms of the RF data as [11], [14], [15]

$$\lambda(\mathbf{g}) = \frac{1}{2}\mathbf{g}^t(\mathbf{K}_0^{-1} - \mathbf{K}_1^{-1})\,\mathbf{g} + \frac{1}{2}\det(\mathbf{K}_1\mathbf{K}_0^{-1}). \tag{5}$$

Ignoring constants and data-independent terms, a sufficient test statistic of the ideal observer can be written more concisely as

$$\lambda_s(\mathbf{g}) = \mathbf{g}^t(\mathbf{K}_0^{-1} - \mathbf{K}_1^{-1})\mathbf{g}. \tag{6}$$

This sufficient test statistic has identical performance to the full log-likelihood. While we use (6) in our paper, the full log-likelihood can be used to take advantage of special properties of log-likelihoods as discussed in [8], [15], and [18]. Note that the test statistic is quadratic in the data. This is to be contrasted with a signal known exactly (SKE) test statistic considered by Zemp *et al.* [13] that is linear in the data. While the linear observer corresponds to a prewhitening matched filter, the quadratic observer (6) corresponds to a prewhitening intensity integrator. Specifically, the strategy of (6) 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 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 **g** is discarded.

D. Monte Carlo Method and Power Series

Abbey [11] introduced a 2AFC Monte Carlo method for assessing the performance of the ideal observer. In this method, pairs of signal present and signal absent stochastic object functions were simulated using random number generators. Signal present and signal absent echo signals were then modeled using the linear systems approach (3). The echo simulation data pairs were then used to compute values for the signal present and signal absent test statistic. The computed test statistic values were then compared to make a decision on which simulated image possessed the lesion. A key contribution was a way of computing the inverse covariance matrices. The covariance matrices are given by $\mathbf{K_g}=\mathbf{H}\mathbf{K_f}\mathbf{H}^t+\sigma_n^2\mathbf{I}$ and are assumed to be known analytically rather than estimated from the Monte Carlo echo data. 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 in the isoplanatic region—thus may be inverted quickly using Fourier techniques. The signal-present covariance is however nonstationary, greatly complicating the inversion. Standard matrix inversion techniques are impractical given the size of the data covariance matrices considered (16 384 × 16 384 for the 128 × 128 images considered below). To accomplish the inversion, a power series was used. For a square matrix **A**

$$(\mathbf{I}-\mathbf{A})^{-1}=\sum_{n=0}^{\infty}\mathbf{A}^n. \tag{7}$$

The series is convergent when the magnitude of the eigenvalues of **A** are less than unity. The inverse signal-present covariance may be written as

$$\mathbf{K}_1^{-1}=\mathbf{K}_0^{-1/2}\left(\mathbf{I}+\mathbf{K}_0^{-1/2}\Delta\mathbf{K}\mathbf{K}_0^{-1/2}\right)^{-1}\mathbf{K}_0^{-1/2}. \tag{8}$$

The inverse operation may be implemented using (7) with $\mathbf{A}=-\mathbf{K}_0^{-1/2}\Delta\mathbf{K}\mathbf{K}_0^{-1/2}$. One does not actually need to deal with these very large covariance matrices to implement the calculation of the test statistic. Instead $\mathbf{K}_0^{-1/2}$ and **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. Importantly, the analytical expressions offer significant intuitive insights that are not obvious from the Monte Carlo approach.

E. 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 [16] of the test statistic, viewed

as the Fourier transform of the probabilities $\mathrm{pr}(\lambda_s|H_1)$ or $\mathrm{pr}(\lambda_s|H_0)$. Consider the characteristic function for the hypothesis H_j (see [17, eqs (17.22)–(17.24a)])

$$\begin{aligned}\psi_j(\xi) &= \int e^{i\xi(\mathbf{g}^t\mathbf{J}\mathbf{g})}\frac{e^{-\frac{1}{2}(\mathbf{g}^t\mathbf{K}_j\mathbf{g}_t)}}{(2\pi)^{NM/2}\sqrt{\det(\mathbf{K}_j)}}d\mathbf{g} \\ &= \frac{1}{\det(I-2i\xi\mathbf{K}_j\mathbf{J})^{1/2}}\end{aligned} \tag{9}$$

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.3 We may also use the identity in [17, eq. (17.13)] to write the denominator of (9) as

$$\begin{aligned}&\det(I-2i\xi\mathbf{K}_1\mathbf{J}) \\ =&\exp\left[\sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}(2i\xi)^m\mathrm{tr}\left[(-\Delta\mathbf{K}\mathbf{K}_0^{-1})^m\right]\right].\end{aligned} \tag{10}$$

The advantage of this form is that the series converges quickly for small contrasts $\Delta\mathbf{K}$. The traces are computationally straightforward to compute by means of the 2-D fast Fourier transform (2DFFT), and they represent the cumulants of the log-likelihood ratio. From a second order truncation of (10), the characteristic function is normal and the signal present distribution density is approximated in the low-contrast limit by a normal distribution (see [17, eqs. (17.26) and (17.27)])

$$\mathrm{pr}_{\mathrm{RF}}(\lambda_s|H_j)\approx N(\mathrm{tr}[\mathbf{K}_j\mathbf{J}],2\mathrm{tr}[(\mathbf{K}_j\mathbf{J})^2]). \tag{11}$$

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 *postwhitened* 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 (11) may be used now to calculate ideal observer detection performance.

F. Detection Signal-to-Noise for SKS Tasks

From the above analysis, the SNR of the sufficient test statistic λ_s may be calculated from (1) and (11) as

$$\mathrm{SNR}^2_{\mathrm{Ideal}}=\frac{\{\mathrm{tr}[\mathbf{J}\mathbf{K}_1]-\mathrm{tr}[\mathbf{J}\mathbf{K}_0]\}^2}{\{\mathrm{tr}[[\mathbf{J}\mathbf{K}_1]^2]+\mathrm{tr}[[\mathbf{J}\mathbf{K}_0]^2]\}}. \tag{12}$$

When pixels are nearly independent, covariance matrices are approximately diagonal and (12) reduces to [4, eq. (33)]. Our approach allows for a more general correlation structure. The power series (7) and (8) can be used to evaluate the traces. A second-order approximation to the ideal observer SNR is

$$\mathrm{SNR}^2_{\substack{\mathrm{Ideal}\\ \mathrm{Low\ Contrast}}}\approx\mathrm{tr}(\mathbf{K}_0^{-1}\Delta\mathbf{K}\mathbf{K}_0^{-1}\Delta\mathbf{K}). \tag{13}$$

[3]For a perfect system, the matrices $\mathbf{K}$ are diagonal, and given enough pixels in the lesion area, the test statistic is a χ^2 random variable. In this case, the characteristic function reduces to a product of χ^2 characteristic functions.

When the test statistic is not normally distributed (e.g., the case of small lesions with few speckle spots), the expression (13) still holds under a low-contrast approximation. In this case, however, although it is a representative performance metric, it may not rigorously parameterize the ROC curve. For higher contrasts, more terms may be retained in the series expansion. Clarkson and Barrett introduced another SNR (5.61 of [15]) 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 [18] converges to (13). Also note that if we consider the full log-likelihood ratio rather than the sufficient test statistic (6), special properties apply [8]. In particular, normal log-likelihood ratios have only one independent moment, say the mean under the signal-present hypothesis [8], and the squared SNR in (12) and (13) is exactly twice this mean (see Barrett and Myers [18, Section 13.2.7] and note that $G(\beta)$ = constant for the Gaussian case. In the next sections, we focus on ways of analytically and numerically evaluating this important expression, and we explain how it can be used for system optimization.

G. White Gaussian Noise (WGN) Object Model

If a WGN object model is assumed for the signal and the background, the signal-present object covariance may be written as $\mathbf{K}_f=\sigma_f^2(\mathbf{I}+\mathbf{S})$, where $\mathbf{S}$ is a diagonal matrix with elements rastorized from the $N\times M$ signal variance template

$$s(\mathbf{x})=\begin{cases}\Delta\sigma^2/\sigma_f^2, & \mathbf{x}\text{ inside lesion}\\ 0, & \mathbf{x}\text{, outside lesion}\end{cases}. \tag{14}$$

The nonzero elements S_i of $\mathbf{S}$, thus, represent a patch of differing variance in the object (see Fig. 1), and are negative for hypo-echoic lesions. They are related to the object contrast factor (OCF) for ultrasound defined by Insana and Hall [24] as the relative change of *standard deviation*: $\text{OCF}=\sqrt{S_i+1}-1$. The signal covariance matrix may be written as $\mathbf{\Delta K}=\mathbf{HSH}^t$ and the covariance of the signal present matrix can, thus, be written as

$$\mathbf{K}_1=\sigma_f^2\mathbf{H}(\mathbf{I}+\mathbf{S})\mathbf{H}^t+\sigma_n^2\mathbf{I}. \tag{15}$$

H. 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 can be used to design systems for maximum task-relevant information content. Truncating the power series expansion, a first-order approximation to the ideal observer test statistic $\lambda=\mathbf{g}^t\mathbf{K}_0^{-1}\mathbf{HSH}^t\mathbf{K}_0^{-1}\mathbf{g}$ (2.9 of [11]) reveals that the ideal observer strategy is to deconvolve the data with $\mathbf{W}\equiv\mathbf{H}^t\mathbf{K}_0^{-1}$ which has the form of a *spatial* Wiener filter.

The filtered echo signal may be envelope detected as is done for B-mode systems. Envelope detection accomplishes two purposes. First, it removes carrier modulation frequencies from the signal. Second, it transforms the zero-mean echo signal into a form where the intensity of the envelope signal is approximately proportional to the object function variance. 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 [11]. Recent coded excitation systems have successfully used deconvolution processing to gain significant advantages in the penetration

of ultrasound systems [19], [20]. Our *detector-level* framework adds the potential to optimize penetration-resolution tradeoffs.

I. Fourier Techniques

The test statistic (6) 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 $\mathbf{K}_0^{-1}=\mathbf{F}^{-1}\mathbf{D}_{K_o^{-1}}\mathbf{F}$ and $\Delta\mathbf{K}=\mathbf{F}^{-1}\mathbf{\Phi}_{\Delta K}\mathbf{F}$. Because $\mathbf{K}_0^{-1}$ is the covariance matrix of a stationary random process it is block-Toeplitz and can be approximated as block circulant [10]. The discrete Karhunen-Loeve transformation is simply the 2-D discrete Fourier transform $\mathbf{F}$, hence, the diagonal matrix $\mathbf{D}_{K_o^{-1}}$ has elements equal to the power spectrum of the background and noise processes $|H(\mathbf{u})|^2 S_b(\mathbf{u}) + S_n(\mathbf{u})$ 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 $\mathbf{\Phi}_{\Delta K}=\mathbf{D}_H\mathbf{\Phi}_S\mathbf{D}_H^*$, where $\mathbf{\Phi}_S=\mathbf{FSF}^{-1}$ is block-circulant and Hermitian, and $\mathbf{D}_H=\mathbf{FHF}^{-1}$ is the diagonal matrix of the eigenvalues of $\mathbf{H}$. 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 (6) for a Monte Carlo study. Its Fourier transform is $(\mathbf{Fg})^t\mathbf{D}_{K_o^{-1}}\mathbf{D}_H\mathbf{FSF}^{-1}\mathbf{D}_H^*\mathbf{D}_{K_o^{-1}}\mathbf{Fg}$. To implement this, we would take the 2DFFT of the RF echo field $\mathbf{g}$ (arranged in the form of an $N\times M$ *matrix*), multiply by the $N\times M$ transfer function $H^*(\mathbf{u})/(|H(\mathbf{u})|^2 S_b(\mathbf{u}) + S_n(\mathbf{u}))$ (which is like a Wiener filter), take an inverse Fourier transform, multiply the variance template $s(\mathbf{x})$ element-by-element, take another 2DFFT, multiply by the transfer function $H^*(\mathbf{u})/(|H(\mathbf{u})|^2 S_b(\mathbf{u}) + S_n(\mathbf{u}))$, then multiply by the Fourier transform of $\mathbf{g}$ 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.

J. 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, (13) becomes

$$\mathrm{SNR}^2_{\substack{\text{Ideal}\\ \text{Low Contrast}}} \approx \mathrm{tr}[\mathbf{D}_{\mathrm{GNEQ}}\mathbf{\Phi}_S\mathbf{D}_{\mathrm{GNEQ}}\mathbf{\Phi}_S] \tag{16}$$

where $\mathbf{D}_{\mathrm{GNEQ}}=\mathbf{D}_H^*\mathbf{D}_{K_o^{-1}}\mathbf{D}_H$. The elements of this diagonal matrix are the elements of the 2-D GNEQ, or *Generalized Noise Equivalent Quanta* [21]

$$\mathrm{GNEQ}(\mathbf{u})=\frac{|H(\mathbf{u})^2|}{|H(\mathbf{u})|^2 S_b(\mathbf{u})+S_n(\mathbf{u})} \tag{17}$$

rastorized into a vector. The *Noise Equivalent Quanta* used in photon modalities is a function of spatial frequency $\mathbf{u}$ and given as $\mathrm{NEQ}(\mathbf{u}) = G^2 \times \mathrm{MTF}(\mathbf{u})/\mathrm{NPS}(\mathbf{u})$, 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 [22] and others [23]. In photon imaging, 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. The concept can be generalized for any imaging technique by considering energy or information density in place of photon density. *Generalized* NEQ as described by Barrett and colleagues [21] 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 $\mathbf{\Phi}_S$ is Hermitian, the trace may be written as $\sum_{\alpha,\beta=1}^{NM}(\mathbf{D}_{\mathrm{GNEQ}})_{\alpha\alpha}|(\mathbf{\Phi}_S)_{\alpha\beta}|^2(\mathbf{D}_{\mathrm{GNEQ}})_{\beta\beta}$. Noting that $\mathbf{\Phi}_S$ is a block-circulant convolution operator, we write this as a continuous integral over spatial frequencies

$$\mathrm{SNR}^2_{\substack{\mathrm{Ideal}\\ \mathrm{Low\ Contrast}}} \approx \int d\mathbf{u}\,\mathrm{GNEQ}(\mathbf{u})[T(\mathbf{u}) \underset{\mathrm{2D}}{*} \mathrm{GNEQ}(\mathbf{u})] \tag{18}$$

where * represents 2-D convolution over spatial frequencies **u**. The integral is evaluated out to the Nyquist limit. 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

$$T(\mathbf{u})=|\mathrm{FT}_{\mathrm{2D}}\{s(\mathbf{x})\}|^2 \tag{19}$$

where $s(\mathbf{x})$ is given by (15). 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 (18) 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

$$\begin{aligned}&\mathrm{SNR}^2_{\substack{\mathrm{Ideal}\\ \mathrm{Low\ Contrast}\\ \mathrm{Large\ Area}}}\\ &\approx \mathrm{AC}^2\int\left[\frac{\mathrm{MTF}^2(\mathbf{u})}{\mathrm{MTF}^2(\mathbf{u})+S_n(\mathbf{u})/|H|_o^2 S_b(\mathbf{u})}\right]^2 d\mathbf{u}\end{aligned} \tag{20}$$

where the modulation transfer function is defined as

$$\mathrm{MTF}(\mathbf{u})=|H(\mathbf{u})|/|H|_o \tag{21}$$

with $|H|_o$ defined as the maximum of $|H(\mathbf{u})|$. C is the lesion contrast given by $C=\Delta\sigma^2/\sigma^2_{\mathrm{obj}}=(1+\mathrm{OCF})^2-1$. The range of validity for the large area approximation will be tested in the numerical results section. The integral in (20) is evaluated over the spatial frequency domain out to the Nyquist limit. Importantly, the integral in (20) is a target-independent figure of merit for characterizing detection performance of ultrasound systems. When noise and object power spectra are flat across the system bandwidth $S_n(\mathbf{u})/|H|_o^2 S_b(\mathbf{u})$ represents the reciprocal of the spatial frequency peak of echo SNR (eSNR), defined as peak signal power spectral density over mean noise power spectral density.

Equation (20) has a form similar to the Smith-Wagner ideal observer SNR which can be written as [4], [24]

$$\mathrm{SNR}^2_{\mathrm{Smith-Wagner}}=\frac{A}{S_c}\mathrm{SNR}_o^2 C^2 \tag{22}$$

where S_c is the correlation area, $\mathrm{SNR}_o^2=1.91$ for Rayleigh statistics of the envelope signal [2], and C is the small-amplitude 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 (20) essentially replaces SNR_o^2/S_c in the Smith-Wagner theory. The new factor includes electronic noise and includes substantial information about the system as encoded in the MTF. Phase information such as phase modulation of the point-spread functions may contribute significantly to the detection performance. The integrand of (20) represents the power spectral density of an ideal noiseless echo signal that is prewhitened in a way that assumes noise is present. The integral is very much like an inverse correlation length of the prewhitened data—or in other words the speckle spot density after prewhitening—thus it reflects the prewhitening potential.

In a previous paper [13], we introduced an SNR metric for ultrasound detection tasks which differs from the description here. In that paper, we considered first a SKE detection task—where the signal was an additive contribution to a background random process. The detection SNR^2 was then averaged over object signal realizations. In that framework the GNEQ also played an important role. An important difference between the SKS theory developed here and SKE theories in previous literature [13], [23] is that there is a quadratic dependence of the GNEQ rather than a linear one. Thus, unlike our previous model [13], the SKS theory predicts that the detection SNR^2 is proportional to the square of the contrast, which is more consistent with the Smith-Wagner description.

So far in this paper we have outlined three main approaches for computing the SKS ideal observer performance: 1) the Monte Carlo approach, (12), using the iterative power series as discussed in [11]; 2) the low-contrast approximation, (18); and 3) the large area, low-contrast approximation, (20). All three have 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 more accurate for large contrasts because the power series is not prematurely truncated, however, it requires extensive computation time compared to (18) and (20), and has less analytic intuition. In Section IV, we wish to show that even for fairly small lesion sizes and moderate lesion contrasts, (20) is a good approximation to the ideal observer. Equations (18) and (20) are the primary theoretical results of this paper.

IV. NUMERICAL RESULTS

A. 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 (18) and (20) with the performance of the Monte Carlo approach using the power series (7) and (8). The Monte Carlo approach is the standard against which our approximations should be measured because the power series is iterated until convergence is achieved. 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 [11]. For each image pair in the 2AFC experiment, the ideal observer test statistic was computed using the power series (10). The power series used up to 16 terms. Contributions of the last term were seen to be below 0.001% of the test statistic sum. For hypo-echoic lesions, the RF image with the smaller test statistic value was scored as the lesion present image (larger test statistic used for hyper-echoic lesions). The percent of correct choices was plotted over a range of contrasts and for three different lesion sizes. The 2AFC SNR is related to percent correct (PC) as $SNR_{2AFC} = 2\Phi^{-1}(PC_{2AFC})$ where Φ is the integral of the standard normal distribution. The YN ideal observer SNR is related to 2AFC SNR by the equation $SNR_{YN}=SNR_{2AFC}/\sqrt{2}$ [5], [24], thus

$$SNR_{YN} = \sqrt{2}\Phi^{-1}(PC_{2AFC}). \quad (23)$$

The results are shown in Fig. 3, (hyper-echoic lesions) and Fig. 4 (hypo-echoic lesions). Note that for hypo-echoic lesions, the large area approximation, (20), over-approximates (18) advantageously for large lesions. Fig. 5 illustrates the representative lesion sizes. Error bars in the ideal observer SNR represent one standard error and correspond to detectabilities evaluated at PC + δPC and PC − δPC, where δPC is the standard binomial sample error $\delta PC = \sqrt{PC(1 - PC)/N_{MC}}$ and N_{MC} is the number of Monte Carlo trials [24]. As expected for low contrast, agreement is very good. Divergence is larger at higher contrasts, but for large lesions the SNR is so high 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 (20). Illustrating this, Fig. 6 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 σ_y and 6.25 σ_x). Fig. 7 shows the predicted SNR of the ideal observer with decreasing electronic noise.

B. System Optimization

In high-noise environments, longer pulse lengths are sometimes desirable to improve signal energy and, hence, tissue penetration. Longer pulses, however, degrade spatial resolution.[4] 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 Fig. 8 and Fig. 9, we plot detectability curves for a range of echo SNRs. Fig. 9 compares the Monte Carlo response and (18) and (20). The approximations under-estimate the Monte Carlo results (due to truncation of the power series) although the predicted peak value remains the same. Higher noise levels could correspond to deeper penetration depths, higher frequency excitations, or poor electronic shielding. In high echo SNR situations, short pulses are seen to be superior, however, in high noise environments, longer pulses are advantageous.

V. DISCUSSION

The results of Fig. 3–Fig. 7 suggests that the low-contrast and large area SNR approximations of the ideal observer are reasonable to use—even for relatively small lesions and moderate contrasts. For small lesions or large contrasts, the approximations do not necessarily agree within the error bars of the Monte Carlo results but are still remarkably close. The SNR expressions can be computed very quickly, thus avoiding time consuming Monte Carlo methods. Equations (18) and (20) 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. 7 illustrates the impact of noise on the ideal observer. Electronic noise hinders the ability of the ideal observer to prewhiten the data. In an ideal noiseless system, the ideal observer performance is only limited by sampling. Although human observer performance (not investigated here) is likely to saturate and plateau with decreasing electronic noise (since human observers have internal noise and, therefore, cannot prewhiten 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.

[4]This discussion refers to pulses with a time-bandwidth product (TBP) of 1. It does not apply to coded excitation where TBP > 1.

For signals sampled using the Nyquist sampling criterion, any increase in sampling rate will have negligible effect on the ideal observer5. Although more noise could potentially be introduced with greater sampling frequencies, the ideal observer optimally filters the noise power to maximize diagnostic detection performance.

A good system design would bandpass filter—thus not integrate all the noise out to the Nyquist limit. Aside from antialiasing filters, we view such bandpass filters as part of processing-level development. Since we wish to focus on detector-level development we neglect such filters. The echo SNR metric we adopt (eSNR) reflects peak power spectral density (PSD) over mean noise PSD. We could have also used total signal power divided by noise power (SNR_{pwr}). Without filtering, the SNR_{pwr} at the optimum point is in Fig. 9 is −5.1 dB, which means that there is nearly twice as much noise power as signal power! However, there is a frequency band where signal PSD is greater than noise PSD (a maximum of 7 dB) and the ideal observer chooses to use these channels to make its decision, and filter to suppress noise. Thus, the image formed by filtering the noise in the optimal way will have SNR_{pwr} significantly above −5.1 dB. Consequently even with this high noise environment, the ideal observer scores more than 80% of its decisions correctly in a 2AFC experiment. We, thus, judged echo SNR based on PSD more appropriate and representative summary measure of detection than SNR_{pwr}. Moreover, when noise and object power-spectral densities are approximately constant over the system bandwidth, 1/eSNR occurs naturally in the denominator of (20). The complete frequency-dependent signal and noise power spectral densities are needed to give a more complete description of the detection task when this is not the case.

In the Smith-Wagner theory shorter pulses (larger band-widths) are always better, since in their derivations, electronic noise was regarded as being less important than speckle and, thus, ignored. Fig. 8 and Fig. 9 illustrate that the optimization point for pulse length is noise dependent. Fig. 9 shows that the Monte Carlo response and (18) and (20) follow the same trend, even though the approximations under-estimate the Monte Carlo results (due to truncation of the power series). Note that from (20) neither lesion contrast nor area will influence the optimum point as long as the assumptions of the theory are met.

In high noise conditions (e.g., high center frequency or deep tissues), there is a tradeoff between penetration and spatial resolution. Longer pulses yield higher echo SNR than short pulses but provide inferior spatial resolution. Fig. 8 shows the optimal tradeoff for detection tasks. It shows that for high echo SNR situations shorter pulse lengths are always better, however as echo SNR decreases, the optimization point shifts toward longer pulses. We see that the system performs moderately well even in high noise conditions. 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. Also, in some cases it may not be cost effective to designing systems with very large ideal observer SNR values: for SNRs above 4 or 5 the gains in percent correct or area under the ROC curve are small (within the measurement uncertainties). The gains in true positive identifications depend on where a radiologist chooses to operate on the ROC curve. Lesion size and contrast will affect the performance levels but will not change the optimization point.

One way of improving spatial resolution could be to use higher transmit frequencies. Large time-bandwidth codes [19], [20], [25] could be used to boost echo SNR to compensate for increased attenuation—thus decreased penetration. This theory tells us how far we can push

[5]This discussion pertains to systems with negligible quantization noise. When there are inadequate quantization levels, oversampling may help to reduce quantization noise, thus contribute to improved detection performance.

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. Fig. 9 shows that even for a 2.95-mm lesion (only 9 pulse lengths) the approximations to the ideal observer detection performance are still reasonable. This means the theory is very applicable to identification of masses below 0.5 cm in diameter, which is the approximate size for which mammography systems are ineffective.

A. Limitations and Extensions

It is important to discuss the limitations of our model. For example, we have neglected aberration which can degrade detection and prewhitening performance. A criterion for a good system design could be the relative robustness to aberration artefacts. 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. Three-dimensional beam properties are a limiting factor when partial volume effects due to elevational clutter are important.

The model has yet to include non-Gaussian statistics or background texture. Inhomogeneous background variability degrades detection and, thus, influences system optimization [26], [27]. Alternative observers could be studied that integrate measured power spectra, similar to techniques discussed by Burgess [27] for mammography. To computationally evaluate the ideal observer test statistic, unknown likelihoods must be estimated. Markov-Chain Monte Carlo techniques [28] have been used for photon imaging modalities to evaluate the test statistic when the background inhomogeneity, signal shape, and location uncertainty can be modeled statistically, and may be applicable to ultrasound.

The detection theory developed here may stand without the need for including background heterogeneity if we assume that clinical images include patches where the detection task approximately reduces to the one that we have modeled, namely, detecting a region of differing echogenicity. Thus, one design strategy could be to optimize the system for low-contrast detectability assuming the worst expected background echo SNR.

A system and processing algorithms should ultimately be designed so that human observer performance is as high as possible. This paper says nothing about modeling human observer performance—a subject of future work. Models of human performance may greatly accelerate processing-level development, since human observer studies can be minimized. It is always a good strategy to optimize ideal observer performance, however it may not be cost effective if human observers cannot access the additional information. Human observer models would allow one to check whether a system design strategy (to optimize the ideal observer performance) and a processing or reconstruction algorithm actually give the desired improvement in human performance. Both human and ideal observer models together should give engineers the tools for task-based system and algorithm design. Finally, any design should be experimentally evaluated *in vitro* on phantoms, and ultimately *in vivo*.

An important practical extension of this work is system optimization for breast cancer imaging; specifically, discrimination between fluid filled cysts and solid masses. For a cyst, the data covariance can be modeled as $\mathbf{K}_{\text{cyst}}=\sigma_f^2\mathbf{H}(\mathbf{I}-\mathbf{S}_{\text{cyst}})\mathbf{H}^t+\sigma_n^2\mathbf{I}$, where $\mathbf{S}_{\text{cyst}}$ is a diagonal matrix consisting of ones for locii inside the cyst, and zeros otherwise. Similarly $\mathbf{K}_{\text{mass}}=\mathbf{K}_1$ defined above for a lesion. The ideal observer test statistic can be computed as $\lambda=\mathbf{g}^t(\mathbf{K}_{\text{cyst}}^{-1}-\mathbf{K}_{\text{mass}}^{-1})\mathbf{g}$. The inversion $\mathbf{K}_{\text{cyst}}$ of and perhaps $\mathbf{K}_{\text{mass}}$ may require a significant number of terms in the power series expansion (7) because of high contrast and will

converge faster in high clutter or high noise environments. Although low-contrast approximations may be inappropriate the performance of this task can be assessed using Monte Carlo techniques. Future work should aim to investigate differences between optimization strategies for detection and discrimination tasks.

B. Application to System Design

Instead of guessing how to trade-off system parameters optimally for detection tasks, we now have a rigorous yet straightforward theory to guide the design. System design specifications can and should 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 [25] 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. With the recent lift of these regulatory limits [29], the combination of *codes* and higher transmit amplitudes may offer significant gains. One may be able to use higher transmit frequencies or harmonic imaging [30], [31] to improve spatial resolution while maintaining echo SNR with codes. Our theory gives a framework for pushing 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 for cyst-mass discrimination tasks [32]. 3) Novel transmit and receive beamforming schemes and lateral codes [13] may also be an open area for additional improvement. 4) Compounding with pre-envelope deconvolved B-mode images is another way of reducing speckle variance while maintaining spatial resolution [33]. This will have the effect of reducing speckle and averaging other artefacts. 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.

VI. CONCLUSION

This paper provides a means for evaluating the ideal observer performance for ultrasonic detection tasks. The approach allows a system to be designed for detection and discrimination tasks based on maximization of diagnostic information in the RF data. Insightful analytical expressions are developed that relate the system RF point-spread function, echo signal-to-noise, lesion contrast, and lesion area to detection performance of the ideal observer. The analytical model describes lesion detection performance of diagnostic systems for hypo- and hyper-echoic lesions as small as only a few speckle spot lengths in diameter. For example, for the 7.2-MHz clinical system in our study, the theory applies to lesions as small as 3 mm and contrasts as large as ~30%. Monte Carlo computational methods are capable of investigating even higher contrasts, and methods may be extended to optimize for cyst-mass discrimination tasks and other related tasks relevant to breast cancer detection, diagnosis, and staging.

Connections are made to *Noise Equivalent Quanta* (NEQ), a detection figure of merit widely 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 defined by the variance profile of the lesion, or its associated Fourier magnitude.

Optimization of engineering tradeoffs such as resolution-penetration can 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 in development, thus 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, conventional B-mode processing suffers an inherent loss of information and 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.

Acknowledgments

The authors are grateful to the Ultrasound Group of Siemens Medical Systems for supplying the Elegra system.

This work was supported in part by the National Institutes of Health (NIH) under Grant RO1 CA 82497. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was E. Krupinski. *Asterisk indicates corresponding author.*

REFERENCES

1. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27 825 patient evaluations. Radiology. 2002 Oct.; vol. 225(no. 1):165–175. [PubMed: 12355001]
2. Wagner RF, Smith SW, Sandrik JM, Lopez H. Statistics of speckle in ultrasound B-scans. IEEE Trans. Sonics Ultrason. 1983 May.vol. SU-30:156–163.
3. Wagner RF, Insana MF, Brown DG. Statistical properties of radio-frequency and envelope-detected signals with applications to medical ultrasound. J. Opt. Soc. Am. A. 1987 May.vol. 4:910–922. [PubMed: 3298583]
4. Smith SW, Wagner RF, Sandrik JM, Lopez H. Low contrast detectability and contrast/detail analysis in medical ultrasound. IEEE Trans. Sonics Ultrason. 1983 May; vol. SU-30(no. 3):164–173.
5. Green, DM.; Swets, JA. Signal Detection Theory and Psychophysics. Los Altos, CA: Penninsula; 1988.
6. Wagner RF. Low contrast sensitivity of radiologic, CT, nuclear medicine, and ultrasound medical imaging systems. IEEE Trans. Med. Imag. 1983 Sep.; vol. MI-2(no. 3):105–121.
7. Barrett HH, Yao J, Rolland JP, Myers KJ. Model observers for assessment of image quality. Proc. Nat. Acad. Sci. USA. 1993 Nov. 1; vol. 90(no. 21):9758–9765. [PubMed: 8234311]
8. Barrett HH, Abbey CK, Clarkson E. Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood generating functions. J. Opt. Soc. Am. A. 1998 June.vol. 15:1520–1535.
9. Zemp RJ, Abbey CK, Insana MF. Linear system models for ultrasonic imaging: Application to signal statistics. IEEE Ultrason, Ferroelectr. Freq. Contr. 2003 Jun.vol. 50(no. 6)
10. Gonzales, RC.; Woods, RE. Digital Image Processing. Reading, MA: Addison-Wesley; 1993.
11. Abbey CK, Zemp RJ, Insana MF. Pre-envelope deconvolution for increased lesion detection efficiency in ultrasonic imaging. Proc. SPIE. 2003; vol. 5034
12. Stavros AT, Thickman D, Rapp C, Dennis M, Parker SH, Sisney GA. Solid breast nodules: Use of sonography to distinguish between Benign and Malignant lesions. Radiology. 1995; vol. 196:123–124. [PubMed: 7784555]
13. Zemp RJ, Abbey CK, Insana MF. Generalized NEQ for assessment of ultrasound image quality. Proc. SPIE. 2003; vol. 5030:391–402.
14. Zemp, RJ.; Abbey, CK.; Insana, MF. Fundamental performance metrics and optimal processing schemes for ultrasonic imaging. Proc. IEEE Ultrasonics Symp.; Germany, Munich. 2002 Oct.. p. 1769-1772.

15. Clarkson E, Barrett HH. Approximations to ideal-observer performance on signal-detection tasks. Appl. Opt. 2000; vol. 39:1783–1793. [PubMed: 18345075]
16. Papoulis, A.; Pillai, SU. Probability, Random Variables and Stochastic Processes. 4th ed.. New York: McGraw-Hill; 2002.
17. Middleton, D. An Introduction to Statistical Communication Theory. Los Altos, CA: Penninsula Publishing; 1987.
18. Barrett, HH.; Myers, KJ. Foundations of Image Science. New York: Wiley; 2003. ch. 18.
19. Haider B, Lewin P, Thomenius KE. Pulse elongation and deconvolution filtering for medical ultrasonic imaging. IEEE Trans. Ultrason. Ferroelectr. Freq. Contr. 1998 Jan..vol. 45(no. 1)
20. Chiao, RY.; Hao, X. Coded excitation for diagnostic ultrasound: A system developer's perspective. Proc. IEEE Ultrason. Symp.; Honalulu, HI. 2003 Oct.. p. 437-448.
21. Barrett HH, Denny JL, Wagner RF, Myers KJ. Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance. J. Opt. Soc. Am. A. 1995 May.vol. 12:834–852.
22. Shaw R. The equivalent quantum efficiency of the photographic process. J. Photog. Sci. 1963; vol. 11:199–204.
23. Wagner RF, Brown DG. Unified SNR analysis of medical imaging systems. Phys. Med. Biol. 1985; vol. 30:489–518.
24. Insana MF, Hall TJ. Visual detection efficiency in ultrasonic imaging: A framework for objective assessment of image quality. J. Acoust. Soc. Am. 1994 Apr.; vol. 95(no. 4):2081–2090.
25. O'Donnell M. Coded excitation system for improving the penetration of real-time phased array imaging systems. IEEE Trans. Ultrason. Ferroelectr. Freq. Contr. 1992 May; vol. 39(no. 3):341–351.
26. Myers KJ, Rolland JP, Barrett HH, Wagner RF. Aperture optimization for emission imaging: effect of a spatially varying background. J. Opt. Soc. Am. A. 1990; vol. 7(no. 7):1279–1293. [PubMed: 2370590]
27. Burgess A. Image quality, the ideal observer, and human performance of radiologic decision tasks. Acad. Radiol. 1995 June; vol. 2(no. 6):522–526. [PubMed: 9419600]
28. Kupinski MA, Hoppin JW, Clarkson E, Barrett HH. Ideal observer computation using Markov chain Monte Carlo. J. Opt. Soc. Am. A. 2003; vol. 20(no. 3):430–438.
29. AIUM/NEMA Standard for Real-Time Display of Thermal and Mechanical Acoustic Output Indices on Diagnostic Ultrasound Equipment, Revision. 1998.
30. Ward B, Baker AC, Humphrey VF. Nonlinear propagation applied to the improvement of resolution in diagnostic medical ultrasound. J. Acoust. Soc. Am. 1997 Jan.; vol. 101(no. 1):143–154. [PubMed: 9000731]
31. Zemp RJ, Tavakkoli J, Cobbold RSC. Computationally efficient modeling of nonlinear ultrasound propagation. J. Acoust. Soc. Am. 2003 Jan.; vol. 113(no. 1):139–152. [PubMed: 12558254]
32. Fernandez AT, Gammelmark KL, Dahl JJ, Keen CG, Gauss RC, Trahey GE. Synthetic elevation beamforming and image acquisition capabilities using an 8 × 128 1.75D array. IEEE Trans. Ultrason. Ferroelectr. Freq. Contr. 2003 Jan.; vol. 50(no. 1):40–57.
33. Trahey GE, Allison JW, Smith SW, von Ramm O. A quantitative approach to speckle reduction via frequency compounding. Ultrason. Imag. 1986 July; vol. 8(no. 3):151–164.

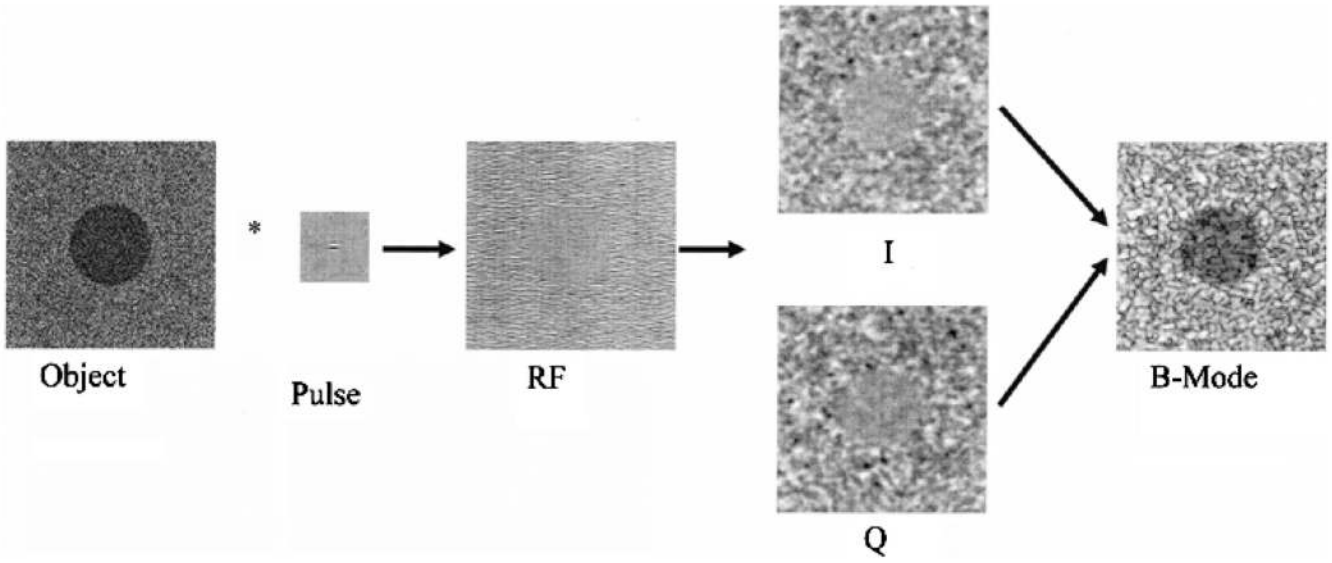

Fig. 1.
Model of the object function magnitude of a circular hypo-echoic lesion, the pulse, RF and IQ data, and the B-mode image.

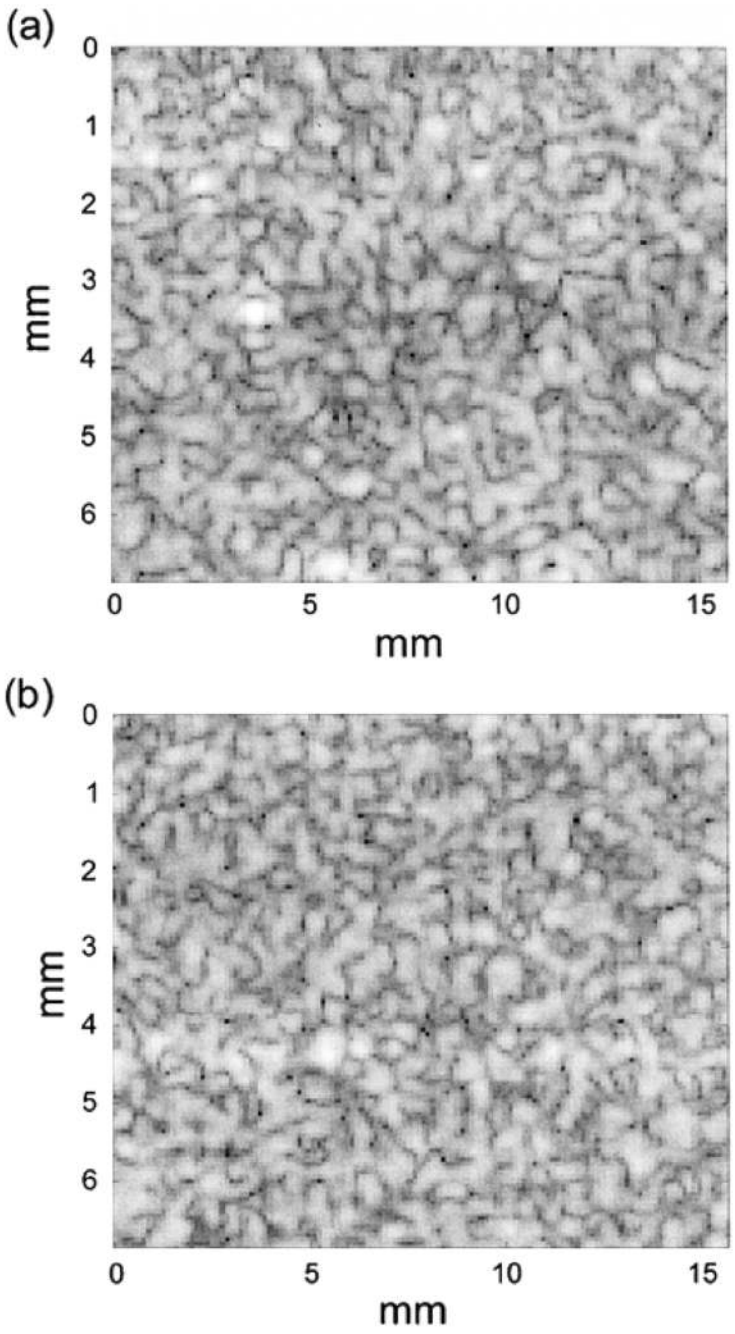

Fig. 2.
(a) B-mode data from a Siemens Elegra system scanning a homogeneous tissue-mimicking phantom (b) Simulated B-mode image.

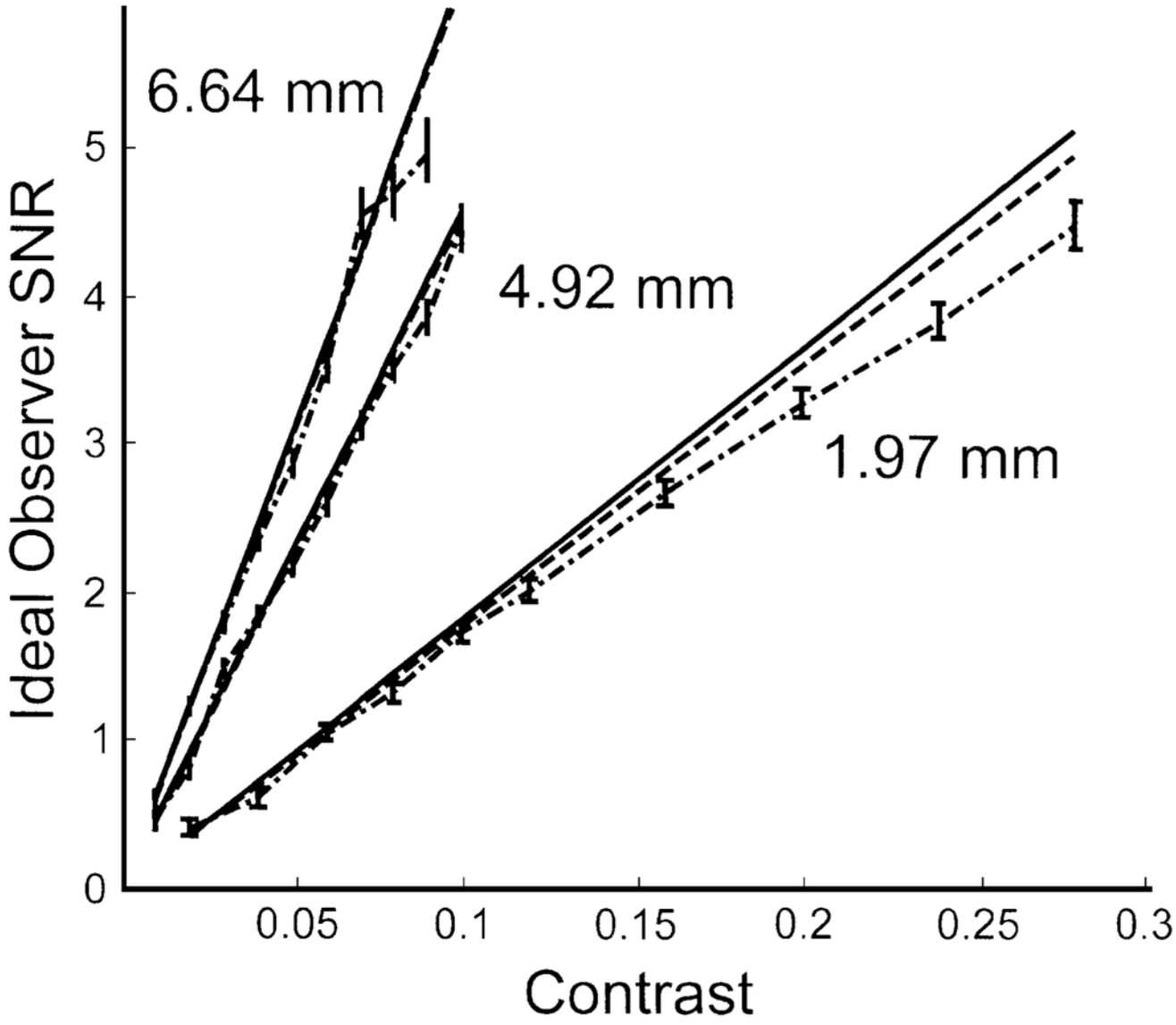

Fig. 3.

A test of the accuracy of the low-contrast approximation over a range of contrasts and for different lesion sizes for *hyper-echoic* lesions. Dot-dashed line: Monte Carlo approach with iterative power series (gold standard), dashed line: (18), solid line: (20). Parameters were otherwise the same as outlined in Section III.

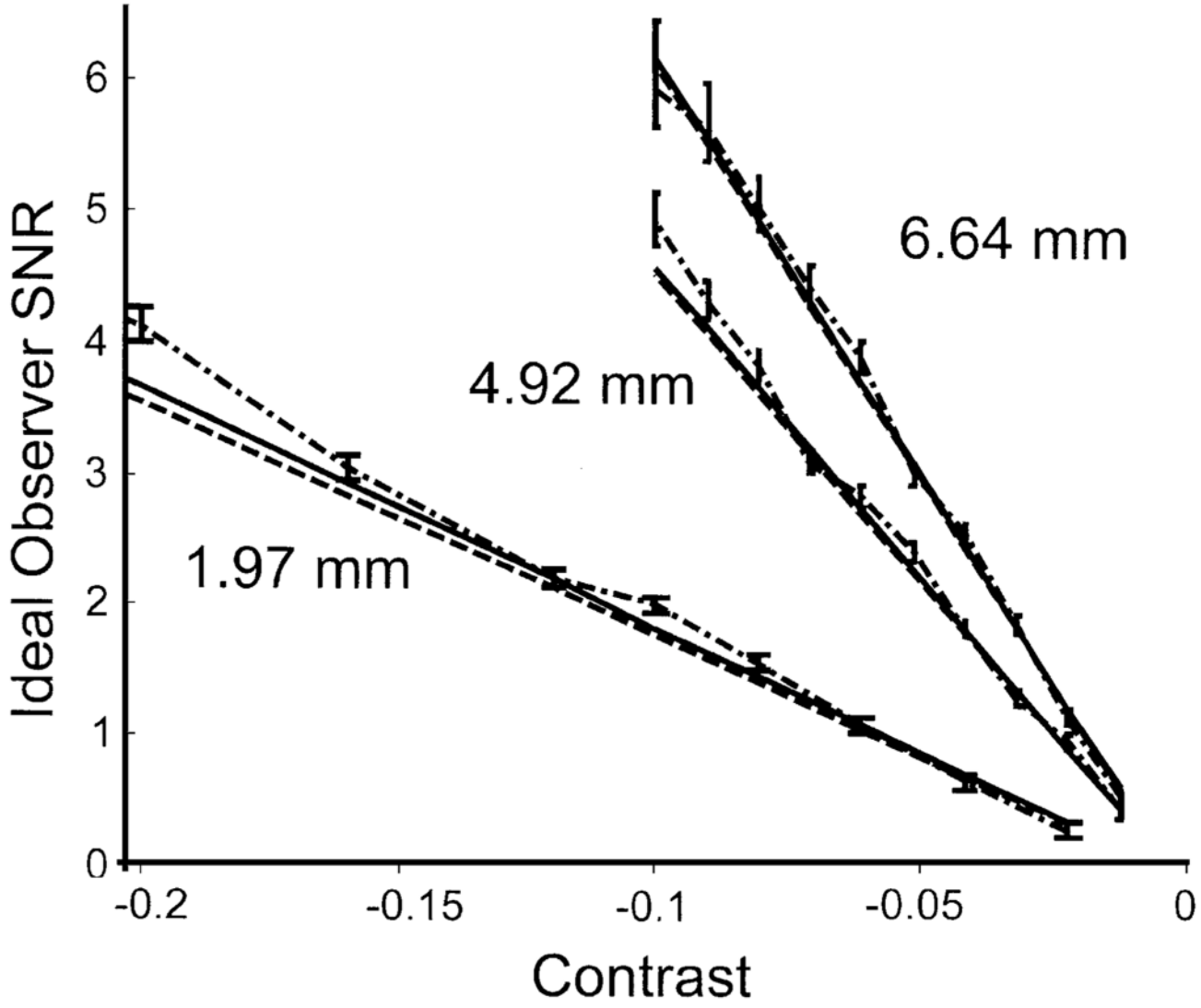

Fig. 4.

A test of the accuracy of the low-contrast approximation over a range of contrasts and for different lesion sizes for *hypo-echoic* lesions. Dot-dashed line: Monte Carlo approach with iterative power series (gold standard), dashed line: (18), solid line: (20). Parameters were otherwise the same as outlined in Section III.

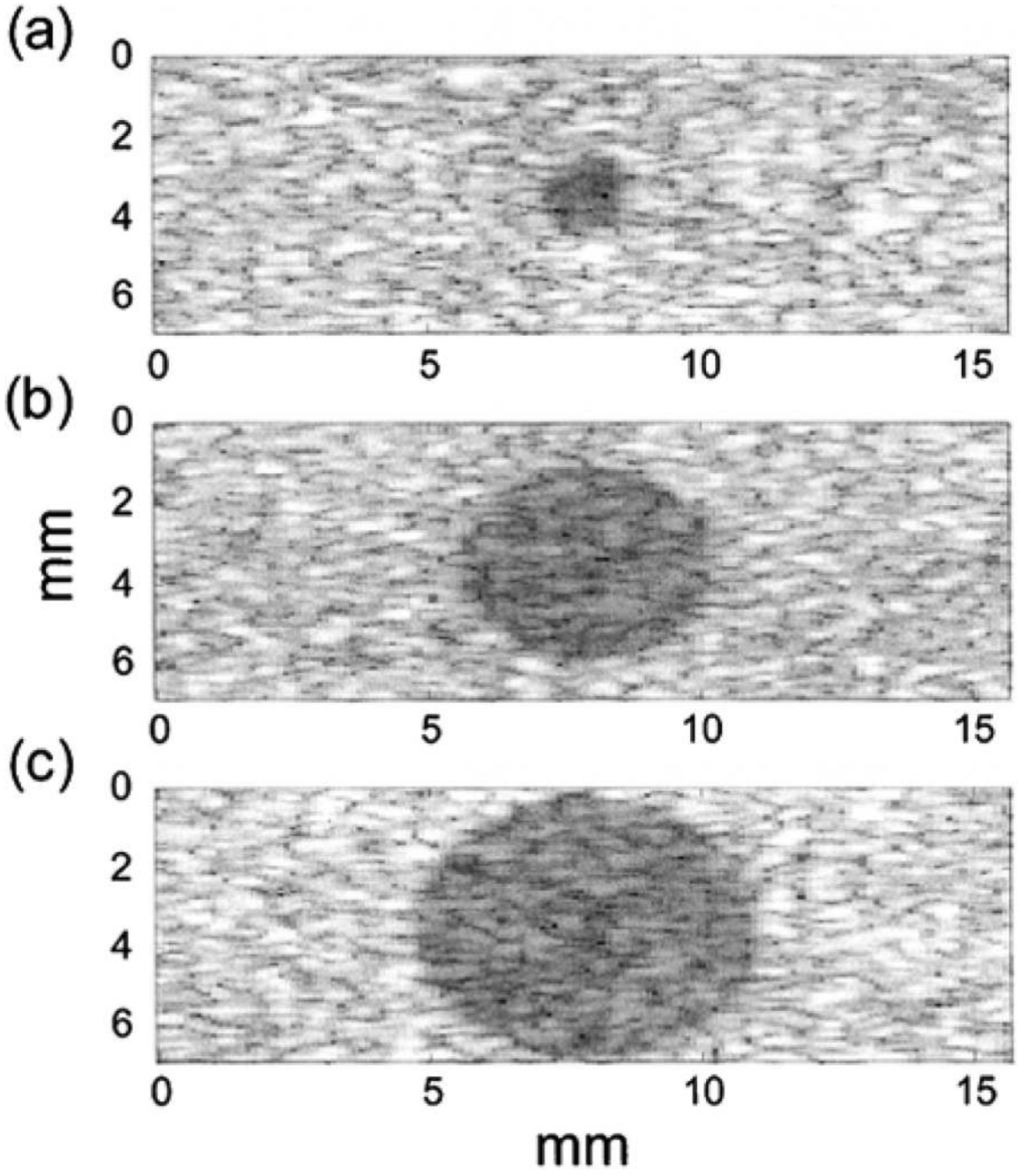

Fig. 5.
Representative lesions sizes used in Fig. 3 and Fig. 4: 1.97-, 4.92-, and 6.64-mm diameters, respectively. Contrast levels were purposely exagerated for visualization.

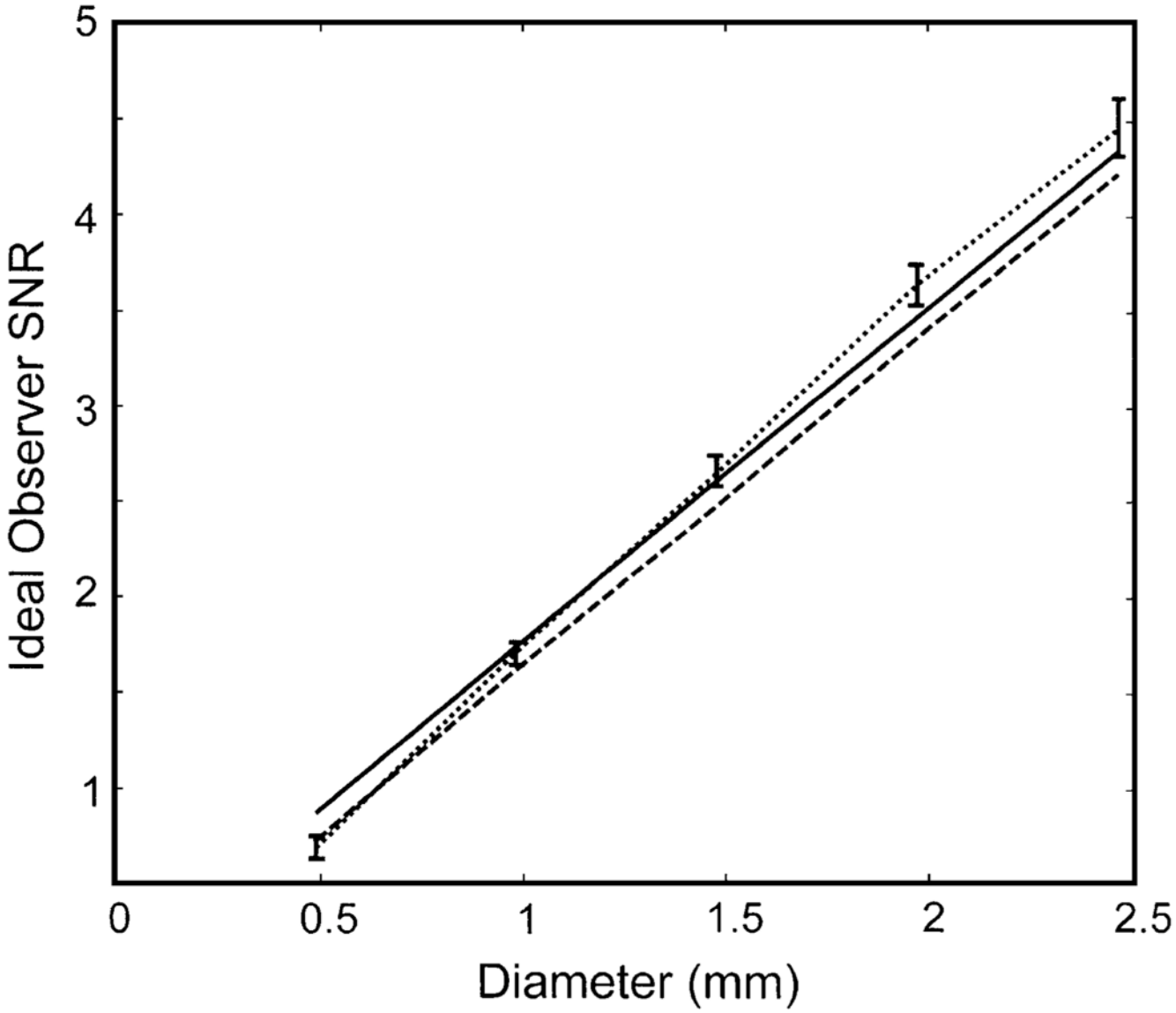

Fig. 6.
A test of the linearity of the ideal observer SNR with lesion diameter for a −19% contrast (hypo-echoic) lesion. Dotted line: Monte Carlo (gold standard), dashed line: (18), Solid line: (20).

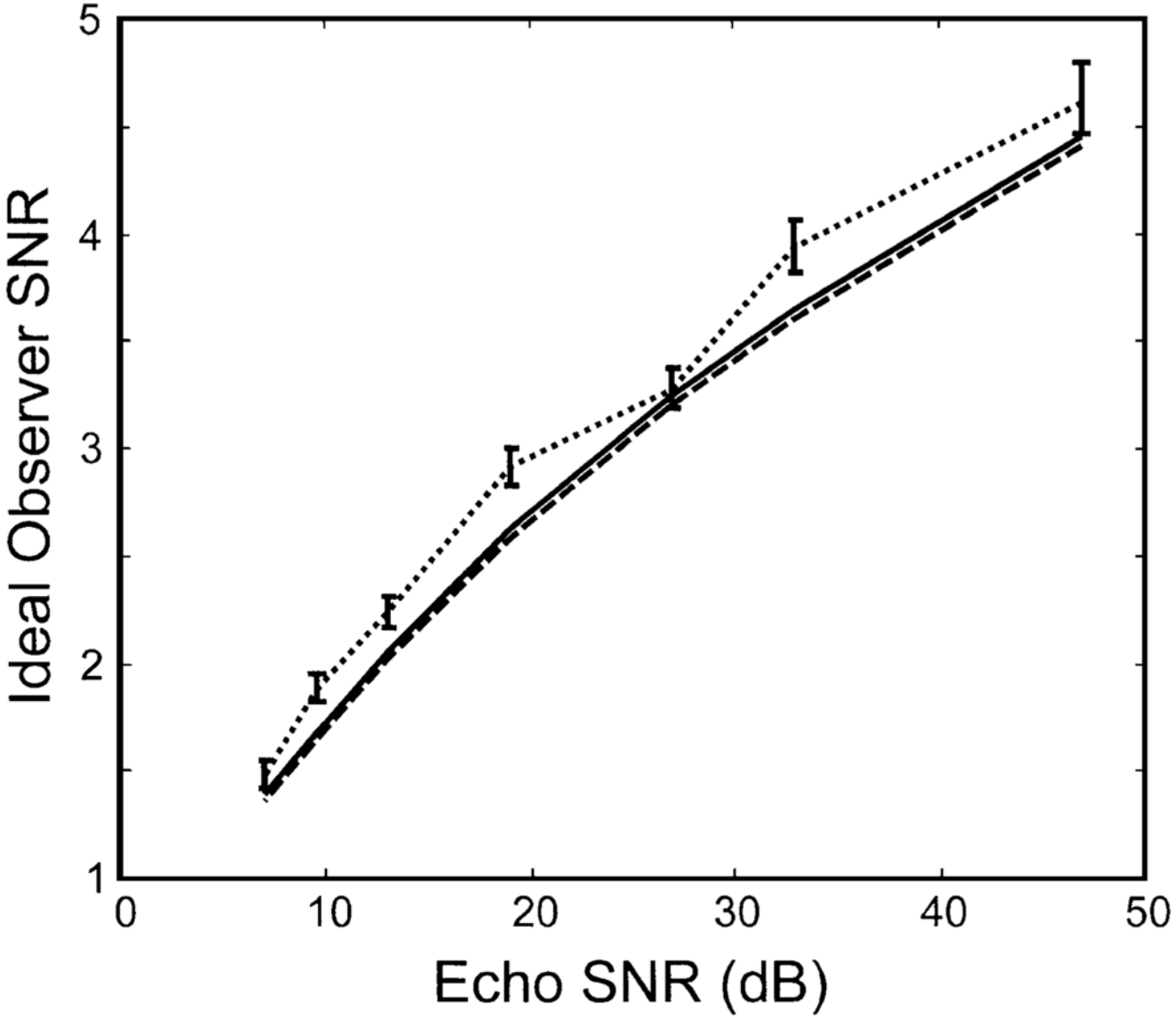

Fig. 7.
Ideal observer SNRs (dotted line: Monte Carlo, dashed line: (18) and solid line: (20)) as a function of varying amounts of electronic noise for a 4.92-mm-diameter lesion of contrast −9.75%. Parameters were otherwise the same as outline in Section III.

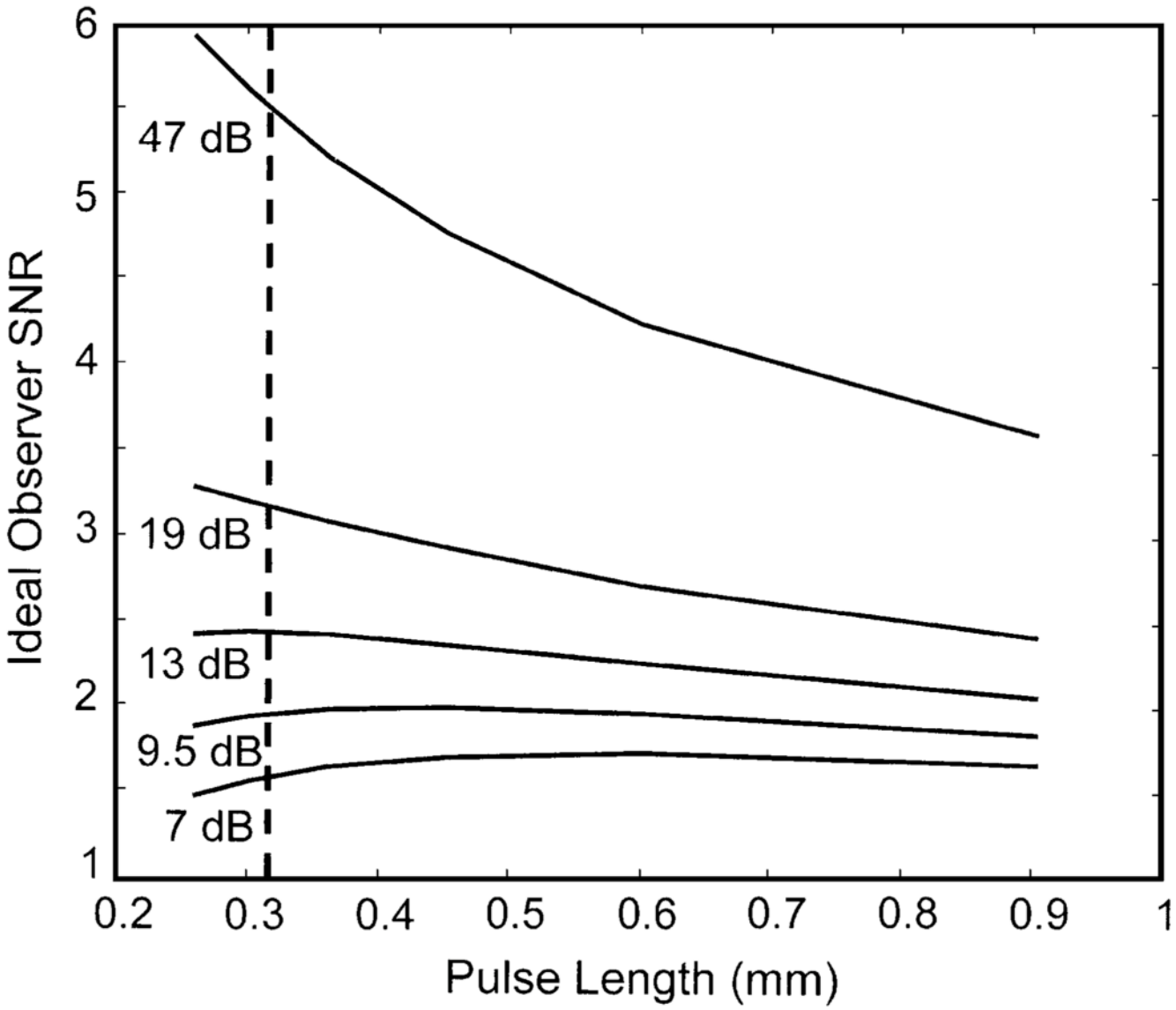

Fig. 8.

An example optimization study: ideal observer SNR as a function of pulse length for a range of echo signal-to-noise levels for a 2.95-mm lesion of contrast of −9.75%. Pulse length is defined as the 20-dB width of the pulse envelope and was measured as 0.316 mm for the Siemens Elegra system at 7.2 MHz. Electronic noise standard deviations of 1, 25, 50, 75, and 100 times that of the Elegra at maximum transmit power were simulated. The echo SNRs will change as a function of pulse length. Echo SNRs of 47, 19, 13, 9.5, and 7 dB existed at these noise levels with a 0.316-mm pulse length (vertical dotted line).

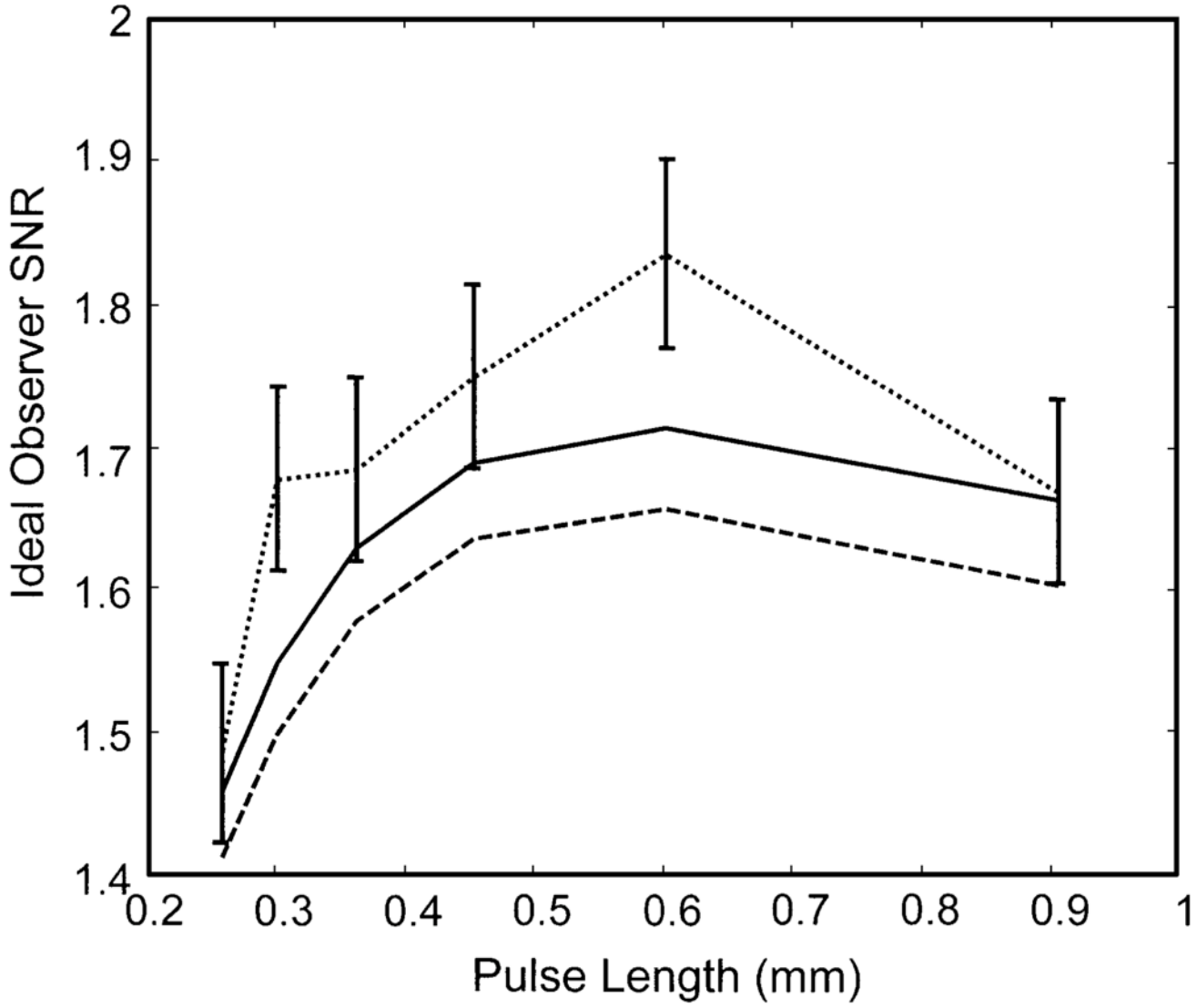

Fig. 9.
Comparison of the Monte Carlo and analytic approximations to Ideal Observer SNR for the 7-dB pulse length optimization in Fig. 8. Dotted Curve: Monte Carlo, solid curve: (20), dashed cuve: (18).