A comparison of material classification techniques for ultrasound inverse imaging

Xiaodong Zhang[a] and Shira L. Broschat[b]
School of Electrical Engineering & Computer Science, Washington State University, P.O. Box 642752, Pullman, Washington 99164-2752

Patrick J. Flynn[c]
Department of Computer Science and Engineering, 384 Fitzpatrick Hall, University of Notre Dame, Notre Dame, Indiana 46556

(Received 2 September 2000; accepted for publication 2 October 2001)

The conjugate gradient method with edge preserving regularization (CGEP) is applied to the ultrasound inverse scattering problem for the early detection of breast tumors. To accelerate image reconstruction, several different pattern classification schemes are introduced into the CGEP algorithm. These classification techniques are compared for a full-sized, two-dimensional breast model. One of these techniques uses two parameters, the sound speed and attenuation, simultaneously to perform classification based on a Bayesian classifier and is called bivariate material classification (BMC). The other two techniques, presented in earlier work, are univariate material classification (UMC) and neural network (NN) classification. BMC is an extension of UMC, the latter using attenuation alone to perform classification, and NN classification uses a neural network. Both noiseless and noisy cases are considered. For the noiseless case, numerical simulations show that the CGEP–BMC method requires 40% fewer iterations than the CGEP method, and the CGEP–NN method requires 55% fewer. The CGEP–BMC and CGEP–NN methods yield more accurate reconstructions than the CGEP method. A quantitative comparison of the CGEP–BMC, CGEP–NN, and GN–UMC methods shows that the CGEP–BMC and CGEP–NN methods are more robust to noise than the GN–UMC method, while all three are similar in computational complexity. © *2002 Acoustical Society of America*. [DOI: 10.1121/1.1424869]

PACS numbers: 43.80.Qf, 43.60.Pt [FD]

I. INTRODUCTION

Breast cancer is a leading cause of women's death in the United States. Each year approximately 182 800 new cases of breast cancer are diagnosed and 40 800 women die.[1] The prognosis for breast cancer is directly correlated with the size of a tumor when detected. The smaller a tumor is when found, the better the chance of survival. Thus it is critical to detect breast cancer when it is at a treatable early stage. This requires a breast screening system with sufficiently high resolution.

X-ray mammography is widely used to detect breast cancer in post-menopausal women. However, for women under the age of 50 x-ray mammography has limited utility because breast tissue in younger women is dense, and dense tissue does not provide the contrast needed to readily image small tumors. This is not the case for ultrasound. Current ultrasound imaging technology is based on a pulse-echo approach in which only the reflected energy is used. However, this works poorly when strong scattering occurs as it does in the premenopausal female breast.[2–4] To account for this strong scattering, full wave inversion techniques can be used.

The full wave inverse problem involves a nonlinear Fredholm integral equation of the first kind which is known to be ill-posed. However, regularization schemes, such as Tikhonov or edge preserving (EP) regularization, can be employed to circumvent the ill-posedness.[5] The known quantities in the integral equation are the measured scattered field and the incident field, while the unknowns are the total field and the object function. The goal of the inverse scattering problem is to reconstruct the object function; the total field is itself a function of the object function. In recent years, much work has been conducted on the inverse problem.[6–18] Most of the algorithms that have been developed are optimization-based techniques, among which the most widely used are the Newton-type methods, the modified gradient (MG) method, and the conjugate gradient (CG) method. Two of the Newton-type methods used are the Gauss–Newton (GN) method[8,9] and the Newton–Kantorovich (NK) method.[11,13] In a comparative study Pichot *et al.* reported that the NK method outperforms the MG method.[17] Although the MG method eliminates the necessity for solving the forward scattering problem at each iteration, it was found to fail when used to reconstruct two large objects simulating the human body. One case was a cylinder with another off-center cylinder inside. The other was an elliptic object with two off-center cylinders inside. In contrast, the NK method was used to successfully reconstruct these two complex objects. Unfortunately, for a different problem with scattered data strongly corrupted by noise (signal-to-noise ratio SNR $=20$ dB or 10% of the maximum amplitude of the scattered data), the NK method diverged.[17] For the same problem, the

[a]Electronic mail: xzhang@eecs.wsu.edu
[b]Electronic mail: shira@eecs.wsu.edu
[c]Electronic mail: flynn@nd.edu

conjugate gradient method with edge preserving regularization (CGEP) converged and yielded satisfactory reconstruction. The CGEP method has also been successfully implemented to reconstruct experimental data.[16] Because of its robustness to noise, the CGEP was chosen for this work.

To reduce the computational cost of the CGEP algorithm, information about the anatomy of the breast and properties of its tissues are incorporated into the algorithm. In earlier work, Manry and Broschat exploited *a priori* information by introducing a material classification technique into the GN method.[9] The addition of material classification improved the rate of convergence of the GN method and also improved the accuracy of the reconstructions. However, the algorithm was found to be sensitive to noise. Motivated by the significant reduction in computational cost provided by the material classifier, a new material classification technique is introduced in this paper. The earlier classifier is termed univariate material classification (UMC) since it uses only values of the tissue attenuation coefficients during the classification process, although information about the sound speeds is also used during a post-classification process. Both classification techniques use a Bayesian classifier, but the new technique utilizes information about the attenuation coefficients and sound speeds of different tissues simultaneously and thus is called bivariate material classification (BMC). The BMC is introduced into the CGEP algorithm after completion of each iteration, and the resulting classified object is used in the next iteration.

In earlier work, a neural network (NN) classification technique was proposed.[18] The NN classifier was used with the CG method and was shown to significantly accelerate convergence of the CG method with better reconstruction accuracy. In this paper, we introduce the NN classifier in the CGEP method. Both the CGEP–NN and CGEP–BMC methods are used to reconstruct a full-sized, two-dimensional breast model. The breast model is the same one that was used in Ref. 9 and is described in the next section. It is composed of layers of skin, fat, and glandular tissue. Tumors and cysts of various sizes are embedded inside the glandular tissue. The model is $32\lambda \times 32\lambda$ in size, where λ is the wavelength in water at a frequency of 400 kHz ($\lambda = 3.72$ mm). It is discretized into 128×128 pixels and is illuminated successively by 128 plane waves at different angles.

Numerical simulations are conducted for both noiseless and noisy cases. By noise we mean that there exists signal noise in the scattered data and that the point-to-point ultrasound parameters of the same tissues vary.[9] Results are presented for a signal-to-noise ratio (SNR) of 70 dB for the scattered field measurements, although noisier cases were also studied. For the noiseless case, the CGEP–BMC correctly reconstructs the object and requires 40% fewer iterations than the CGEP method, while the CGEP–NN method correctly reconstructs the object and requires 55% fewer iterations than the CGEP method. For the noisy case, tumors and cysts are detectable for all three methods, and minimal tumors and a small cyst are correctly classified by the BMC and NN classifiers. However, some pixels are misclassified or are left unclassified. A quantitative comparison of the CGEP–BMC, CGEP–NN, and GN–UMC methods shows that the CGEP–BMC and CGEP–NN methods are more robust to noise than the GN–UMC method, while they all have comparable computational costs.

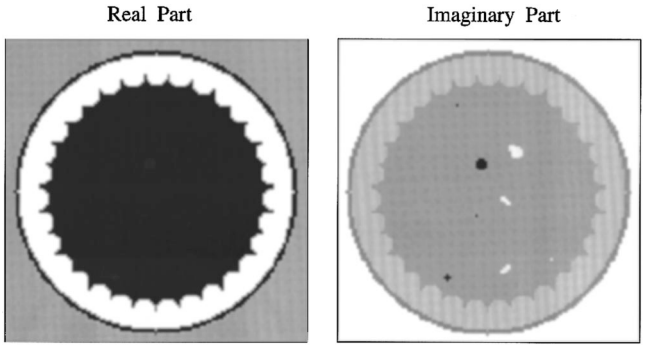

FIG. 1. True contrast for the breast model.

In Sec. II, the two-dimensional breast model is described. The CGEP method is discussed in Sec. III. In Sec. IV, the BMC classifier is presented, and in Sec. V, the neural network classifier is briefly discussed. Numerical results are presented in Sec. VI. Finally, concluding remarks are given in Sec. VII.

II. 2D BREAST MODEL AND SCATTERING GEOMETRY

The two-dimensional breast model used in the numerical simulations is shown in Fig. 1. In the real part of the object the tumors and cysts are difficult to discern, while they stand out in the imaginary part. The real part of the object depends mostly on the speed of sound while the imaginary part depends mostly on the attenuation. The lack of contrast in the real part reflects the similarity of the tissue sound speeds. On the other hand, the attenuations of glandular tissue, cysts, and tumors are distinct, and this is reflected in the imaginary part of the object.

The breast model is the same as that used in Ref. 9 and is discretized into 128×128 pixels, corresponding to an average breast with a 12-centimeter diameter. Simulations are performed at a frequency of 400 kHz with each pixel about 0.93 mm wide ($\lambda/4$ is water). Since inverse imaging is a full wave solution, it has a resolution of approximately $\lambda/4$. Thus, 400 kHz provides a resolution of less than a millimeter in water which should be adequate for breast imaging. The model is comprised of a skin layer 1.86 mm thick with an outer radius of 5.59 cm. Next a fat layer is constructed with an annular layer 5 mm wide and 1-cm fat lobes centered along the inner radius of the annulus. This geometry creates a highly refractive fat-glandular interface which provides a rigorous test of the imaging method. The breast interior is composed of glandular tissue and tumors and cysts of various sizes and shapes. Minimal tumors and a cyst one pixel in size (0.93 mm) are used to test the ability to image small objects. The model is assumed to be surrounded by water and is illuminated by 128 plane waves equally spaced around it. The scattered field is measured for each illumination along the four edges by 128 receivers on each edge. The measured scattered data for the true object are calculated using the

TABLE I. Typical tissue parameters at f=400 kHz.

Tissue	Speed [m/s]	Attenuation [m^{-1}]
Fat	1439.0	1.66
Glandular	1581.0	2.90
Skin	1586.0	3.36
Tumor	1573.0	6.58
Cyst	1584.0	0.46
Water	1490.0	0.0

BI–CGSTAB–FFT method.[19] Tissue parameters for the breast at a frequency of 400 kHz are listed in Table I as presented by Madsen *et al.*[20]

III. THE CONJUGATE GRADIENT METHOD WITH EDGE PRESERVING REGULARIZATION

Figure 2 shows the two-dimensional scattering geometry. The computational domain D is defined to be square. M receivers are located on a surface S enclosing D. The acoustic sound speed at r is denoted by $c(r)$, and the attenuation coefficient is $\alpha(r)$. The sound speed in the surrounding medium (here, water) is c_0, and the surrounding medium is lossless. The object is irradiated successively by plane acoustic waves at L different angles, l=1,2,...,L.

Assuming tissue density variations are negligible, i.e., that breast tissue is generally soft, the acoustic wave equation can be cast into integral equation form[21]

$$u_l(r)=u_l^{\text{inc}}(r)+\int_D u_l(r')o(r')g(r,r')dr', \quad r\in D, \tag{1}$$

$$u_l^s(r)=\int_D u(r')o(r')g(r,r')dr', \quad r\in S \tag{2}$$

for the lth excitation where $u_l(r)$ is the total field, $u_l^{\text{inc}}(r)$ is the incident field, $u_l^s(r)$ is the scattered field, and $g(r,r')$ is the Green's function. The object function $o(r)$ satisfies the relationship $o(r)=k_0^2\{[k^2(r)/k_0^2]-1\}$ where k_0 is the wave number in the background medium (water), and $k(r)$ is the wave number at r with $k(r)=[2\pi f/c(r)]-j\alpha(r)$. The term $\{[k^2(r)/k_0^2]-1\}$ is called the object contrast.

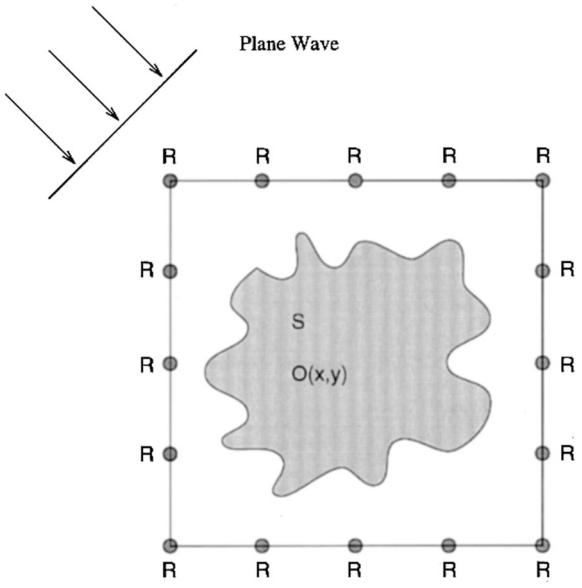

FIG. 2. Scattering geometry.

The goal of inverse imaging is to reconstruct the object function $o(r)$ using knowledge of the scattered data u^s. This problem is ill-posed by nature in the Hadamard sense[5]—i.e, the existence, uniqueness, and stability of the solution are not simultaneously ensured.[22] Using the moment method with pulse basis functions and point matching,[23] Eqs. (1) and (2) are converted to matrices with the domain D discretized into N square cells. This yields a nonlinear matrix system for each illumination l (l=1,2,...,L):

$$\vec{u}_l=(I-G_DO)^{-1}\vec{u}_l^{\text{inc}}=\mathcal{L}(O)\vec{u}_l^{\text{inc}}, \tag{3}$$

$$\vec{u}_l^s=G_SO\mathcal{L}(O)\vec{u}_l^{\text{inc}}=G_SO\vec{u}_l, \tag{4}$$

where O is an $N\times N$ diagonal matrix with the diagonal elements forming the object vector $\vec{o}$, $\vec{u}_l^s$ is the $M\times 1$ scattered field vector, $\vec{u}_l$ is the $N\times 1$ total field vector, and $\vec{u}_l^{\text{inc}}$ is the incident field vector. The subscript l indicates the lth illumination, and G_D and G_S are the $N\times N$ and $M\times N$ Green's function matrices, respectively. The elements for G_D and G_S are

$$G_{D_{ij}}=\int_{D_j} g(r_i-r')dr', \quad r\in D, \tag{5}$$

$$G_{S_{ij}}=\int_{D_j} g(r_i-r')dr', \quad r\in S, \tag{6}$$

where D_j is the area of the jth cell, and r_i is the center of the ith cell. G_D maps the domain D to domain D and G_S maps the domain D to domain S.

The inverse scattering problem requires solution of both a forward and an inverse problem. When the object is known, Eq. (3) is used to solve for the total field $\vec{u}_l$; the problem is well-posed and linear with respect to the incident field $\vec{u}_l^{\text{inc}}$. This is the forward problem which is solved using the BI–CGSTAB–FFT method.[19,24] The total field of Eq. (3) is expressed in terms of basis functions, and the unknowns are the coefficients of these basis functions. This linear system is solved using the BI–CGSTAB method while the kernal of the integral operator is computed via FFTs. When the total field is known and Eq. (4) is used to solve for the object $\vec{o}$, the problem is nonlinear and ill-posed. This is the inverse problem. The ill-posedness inherent in Eq. (4) is more difficult to handle because of the nonlinearity.

A. The conjugate gradient technique

Since the CG method is an optimization approach, an error function must be defined

$$\vec{\rho}_l=\vec{u}_l^s-G_SO\mathcal{L}(O)\vec{u}_l^{\text{inc}}, \tag{7}$$

where $\vec{u}_l^s$ is the true scattered data vector ($M\times 1$) for illumination by the lth source, O is an $N\times N$ diagonal matrix whose elements are the elements of the object vector $\vec{o}$ ($N\times 1$), $\vec{u}_l^{\text{inc}}$ is the incident field vector ($N\times 1$), $\mathcal{L}(O)=(I-G_DO)^{-1}$, and the matrices G_D and G_S are defined by Eqs. (5) and (6), respectively. N is the number of cells into which the object is discretized, and M is the num-

ber of receivers per view. In addition, a cost functional is defined

$$\Phi = \frac{1}{2} w \sum_{l=1}^{L} \|\vec{\rho}_l\|^2, \tag{8}$$

where $\|\cdot\|$ is the Euclidean norm. Thus, the functional Φ is a measure of the discrepancy between the measured and calculated scattered data due to estimation of the object. As proposed in Ref. 10, the weight w is defined by

$$w = \left(\sum_{l=1}^{L} \|\vec{u}_l^s\|^2 \right)^{-1}, \tag{9}$$

where summation is over all sources. This choice ensures that the amplitude of the incident field has no impact on the function since $\vec{u}^s$ is linear in $\vec{u}^{\text{inc}}$.

The CG method involves several key updating equations:

$$\hat{\vec{o}}_{n+1} = \hat{\vec{o}}_n + t_n \vec{d}_n, \tag{10}$$

$$\vec{d}_{n+1} = -\vec{g}_{n+1} + \gamma_n \vec{d}_n, \tag{11}$$

where t_n is the step size along $\vec{d}_n$, $\vec{d}_n$ is the updating direction for the object at the nth iteration, $\vec{g}_n$ is the gradient of the cost functional at the nth iteration, and γ_n is the CG direction parameter. $\vec{o}_1$ is the initial guess, and $\vec{d}_1 = -\vec{g}_1$. Thus, we use the steepest descent direction to initiate the iterative process. Choices for t_n and γ_n are discussed below. The gradient of the cost functional is obtained from

$$\vec{g}_n = w \sum_{l=1}^{L} [\text{diag}(\mathcal{L}(\hat{O}_n)\vec{u}_l^{\text{inc}}) \mathcal{L}(\hat{O}_n)]^{\dagger} G_s^{\dagger} \vec{\rho}_{l,n}, \tag{12}$$

where $\mathcal{L}(O) = (I - G_D O)^{-1}$, $\dagger$ indicates the conjugate transpose, diag denotes the diagonal, and $\vec{\rho}_{l,n}$ is the difference between the measured and calculated scattered data at the nth iteration when the object is illuminated by the lth source. $\mathcal{L}(\hat{O}_n)\vec{u}_l^{\text{inc}}$ is essentially the total field at the nth iteration and is calculated for the forward problem.

To save computational time, a first-order approximation for t_n for large objects is used

$$t_n = \frac{\Sigma_{l=1}^{L} \langle \vec{\rho}_l, \vec{V}_l \rangle}{\Sigma_{l=1}^{L} \|\vec{V}_l\|^2}, \tag{13}$$

where $\vec{V}_l = G_s\{\mathcal{L}(\hat{O}_n)\}^T \text{diag}(\vec{d}_n)\mathcal{L}(\hat{O}_n)\vec{u}_l^{\text{inc}}$, ''$T$'' means transpose, and the angle brackets denote the inner product.

There are many choices for the conjugate gradient direction parameter γ_n. In a previous study it was found that the recently proposed direction parameter given by

$$\gamma_n = \frac{\langle \vec{g}_{n+1}, \vec{g}_{n+1} \rangle}{\langle \vec{d}_n, \vec{g}_{n+1} - \vec{g}_n \rangle}$$

is the best choice for ultrasound inverse imaging.[25]

B. Edge preserving regularization

To further enhance the performance of the conjugate gradient method when the problem is strongly ill-posed and/or the scattered data is strongly corrupted, an edge preserving (EP) regularization technique is introduced into the conjugate gradient method. It allows smoothing of a homogeneous region while preserving edges in an image. It is a significant improvement over Tikhonov regularization[26,27] and, in addition, it is convenient to apply with the conjugate gradient method.

The original edge preserving regularization proposed by Lobel *et al.* was imposed on both the real and imaginary parts of the image.[10] For ultrasound imaging of the breast, edge preserving regularization is imposed on just the imaginary part of an image because only the imaginary part gives discernible results (see Fig. 1). Edge preserving regularization has the form

$$\Phi_R = \mu \sum_{p=1}^{N_{\text{lin}}} \sum_{q=1}^{N_{\text{col}}} \left(b_{p,q} \left\| \frac{\text{Im}(\nabla o)_{p,q}}{\delta} \right\|^2 + \Psi(b_{p,q}) \right), \tag{14}$$

where Im indicates the imaginary part. μ is the regularization parameter which balances the data term and the regularization term. The parameter δ fixes the threshold level of the gradient norm above which a discontinuity is preserved and below which it is smoothed. For this project, δ was chosen empirically. The $b_{p,q}$ terms are real and continuous in [0,1], p is the pixel number in the x direction, and q is the pixel number in the y direction. $(\nabla o)_{p,q}$ is the gradient of the object function at position (p,q). The $b_{p,q}$ terms record discontinuities, i.e., edges, in the imaginary part of the image. The function Ψ is given by[26]

$$\Psi(s) = s + \frac{1}{s} - 2. \tag{15}$$

The new cost functional is given by

$$\Phi_{\text{EP}} = w \sum_{l=1}^{L} \|\vec{\rho}_l\|^2 + \mu \sum_{p=1}^{N_{\text{lin}}} \sum_{q=1}^{N_{\text{col}}} \left(b_{p,q} \left\| \frac{\text{Im}(\nabla o)_{p,q}}{\delta} \right\|^2 + \Psi(b_{p,q}) \right) \tag{16}$$

which depends on two variables, $\vec{o}$ and $b_{p,q}$. To minimize Eq. (16), we use successive over-relaxation, alternating minimization of Φ by fixing one of the two variables successively.

(1) When the $b_{p,q}$ terms are fixed, minimization of Eq. (16) is performed using the conjugate gradient method as discussed in the preceding section.
(2) When $\vec{o}$ is fixed, the values for the $b_{p,q}$ terms that minimize Φ_{EP} are unique and are given for each point (p,q) by the analytical expression

$$b_{p,q} = \frac{1}{2\sqrt{1 + \left(\frac{\|\text{Im}(\nabla o)_{p,q}\|}{\delta} \right)^2}}. \tag{17}$$

Since $b_{p,q}$ gives the spatial gradient of the imaginary part of the object, it records information about discontinuities at each step of the algorithm. The values of $b_{p,q}$ from the previous step are used for the next estimate of the object. For simplicity, we assume the operators ∇ and Im are inter-

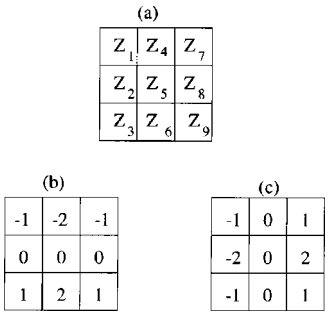

FIG. 3. (a) 3×3 image region, (b) mask used to compute G_y at the center pixel, and (c) mask used to compute G_x at the center pixel. These masks are known as the Sobel operator.

changeable, $\mathrm{Im}(\nabla o)_{p,q}=\nabla(\mathrm{Im}(o_{p,q}))$. To simplify notation, we set $\mathrm{Im}(o_{p,q})=Z_{p,q}$ so that $\mathrm{Im}(\nabla o)_{p,q}=\nabla Z_{p,q}$.

In image processing it is common to approximate $|\nabla Z_{p,q}|$ by[28]

$$|\nabla Z_{p,q}|=|G_{x_{p,q}}|+|G_{y_{p,q}}|, \tag{18}$$

where $G_{x_{p,q}}$ and $G_{y_{p,q}}$ are the gradient operators in the x and y directions, respectively. Sobel operators were chosen to compute the gradients because they provide both differencing and smoothing.[28] Since derivatives tend to enhance noise, smoothing is a particularly attractive feature. A Sobel operator operates on a 3×3 pixel section of an image as shown in Fig. 3 and is also known as a mask. The gradients at the center pixel, denoted Z_5, are given by

$$G_x=(Z_7+2Z_8+Z_9)-(Z_1+2Z_2+Z_3), \tag{19}$$

$$G_y=(Z_3+2Z_6+Z_9)-(Z_1+2Z_4+Z_7). \tag{20}$$

After the gradients for the center pixel have been determined, the masks are moved to the next pixel location and the procedure is repeated. The values along the four edges of the image are set to the values at the corners of the image. For our problem, the gradient of Φ_R with respect to the imaginary part of the contrast at location (p,q) is given by

$$\begin{aligned}\frac{\partial\Phi_R}{\partial Z_{p,q}}=\frac{1}{\delta^2}\{&b_{p-1,q-1}(G_{y_{p-1,q-1}}+G_{x_{p-1,q-1}})\\&+b_{p,q-1}(2G_{y_{p,q-1}})+b_{p+1,q-1}(G_{y_{p+1,q-1}}\\&+G_{x_{p+1,q-1}})+b_{p-1,q}(2G_{x_{p-1,q}})\\&+b_{p+1,q}(-2G_{x_{p+1,q}})+b_{p-1,q+1}(-G_{y_{p-1,q+1}}\\&+G_{x_{p-1,q+1}})+b_{p,q+1}(-2G_{y_{p,q+1}})\\&+b_{p+1,q+1}(-G_{y_{p+1,q+1}}+G_{x_{p+1,q+1}})\},\end{aligned} \tag{21}$$

where

$$\begin{aligned}G_{x_{p,q}}=&(Z_{p+1,q-1}+2Z_{p+1,q}+Z_{p+1,q+1})\\&-(Z_{p-1,q-1}+2Z_{p-1,q}+Z_{p-1,q-1})\end{aligned}$$

and

$$\begin{aligned}G_{y_{p,q}}=&(Z_{p-1,q+1}+2Z_{p,q+1}+Z_{p+1,q+1})\\&-(Z_{p-1,q-1}+2Z_{p,q-1}+Z_{p+1,q-1}).\end{aligned}$$

The steps for the CGEP iteration procedure are

(1) Set initial guess.
(2) Repeat the following until convergence occurs.

(a) Minimize Eq. (16) with respect to $b_{p,q}$ when $\vec{o}$ is fixed. The updating equation for $b_{p,q}$ is given by Eq. (17).
(b) Minimize Eq. (16) with respect to $\vec{o}$ when the $b_{p,q}$ terms are fixed. Use the CG algorithm to estimate $\hat{\vec{o}}_{n+1}$ as discussed earlier.

The gradient of Φ_{EP} must now have two parts, the gradient $\vec{g}_1$ due to Φ and the gradient $\vec{g}_2$ due to Φ_R. Thus, $\vec{g}=\vec{g}_1+\vec{g}_2$. Equation (13) is no longer valid and must be rederived to be completely rigorous. However, Eq. (13) is a reasonable approximation to the actual equation when μ is small. Since $\mu=10^{-6}$ is chosen for this work, Eq. (13) is used. δ^2 is chosen empirically to be 250.

IV. BIVARIATE MATERIAL CLASSIFICATION (BMC)

To accelerate convergence of the CGEP algorithm, a new material classification technique is introduced into the CGEP iterations. It exploits knowledge of the material parameters, specifically the sound speed and attenuation, and thus the name bivariate material classification (BMC). Each pixel (p,q) of the object is represented by its sound speed x and attenuation y, and speed and attenuation are assumed to be both uncorrelated and Gaussian distributed. The BMC classifier is essentially a Bayesian classifier,

$$\wp=\max_c\frac{P(c)p(x,y|c)}{p(x,y)}, \tag{22}$$

where c is the tissue class type (e.g., cyst, fat, skin, tumor), $P(c)$ is the *a priori* probability of class c, $p(x,y|c)$ is the conditional probability density function given a class c, and $p(x,y)=\Sigma_c P(c)p(x,y|c)$ is the total probability density function. For convenience, attenuation values are weighted by a factor of 500. Since $P(c)p(x,y|c)/p(x,y)$ is the *a posteriori* probability $P(c|x,y)$ of a pixel (x,y) belonging to class c, $\wp\in[0,1]$. Thus $\wp$ is the maximum value of the *a posteriori* probability at pixel (x,y) for all classes considered. For example, after computation of the *a posteriori* probabilities for all classes considered, if fat tissue yields the maximum value, then the pixel at (x,y) is classified as fat and $\wp$ is set to be the *a posteriori* probability of fat. Information about the structure of the breast is also utilized. For example, it is known that skin tissue does not exist in the interior of the breast, and this fact is used to test for misclassification by the BMC classifier.

To implement bivariate material classification, we need values for the conditional probability density function (PDF) $p(x,y|c)$ and the probability $P(c)$ for each class. The strategy for finding these values is similar to the one proposed in Ref. 9 except that both the sound speed and attenuation are used. The conditional PDF $p(x,y|c)$ is assumed to be Gaussian, and the sound speed x and attenuation y are assumed to be independent. The mean values of both the sound speed and attenuation for all classes are already known, as shown in Table I. Thus, the task reduces to finding the variance of

the conditional PDF and the probability $P(c)$. The reconstructed object is used to create histograms of the sound speed and attenuation, and these are used to estimate the variance and probability. For each histogram, a rectangular window is centered at the mean value for a particular class, and the width of the window is determined empirically. The probability $P(c)$ is estimated by the percentage of pixels within the window of each class. The variance of the conditional PDF is found in a similar manner.

Once the conditional PDF $p(x,y|c_j)$ and the probability $P(c_j)$ for $j=1,2,...,C$ are obtained, we are ready to perform classification for each pixel. However, there are situations when no decision should be made, for example, when it is equally likely that a pixel is two different tissue types. As proposed by Ref. 9, a "rejection" class is introduced to handle such situations. The rule for determining a rejection is

If $\wp \geq 1-\Re$, set tissue to c, else reject.

where $0 \leq \Re \leq 1$ is the rejection rate, a parameter for controlling classification. If $\Re=0$, then decisions from the Bayesian classifier are rejected; if $\Re=1$, then decisions are accepted. Thus $\Re$ is used as a measure of confidence in the ability to classify. The BMC classifier is used in the CGEP algorithm at every 10th iteration. $\Re$ is increased linearly from 0.5 to 0.9 from the 10th iteration to the 50th iteration. This has been found to give reasonable results.

Since decisions made by the Bayesian classifier can be incorrect, a "clean up" step is inserted after the Bayes classification step to prevent misclassification. The sound speed and attenuation of nonrejected pixels are checked to ensure that they fall within acceptable bounds. These bounds are chosen on the basis of the known ultrasound properties for the possible tissue types. For example, as indicated in Table I, the range of possible sound speeds is 1400 to 1800 m/s. If the bounds are exceeded, pixel identification is rejected. Location of the tissue type is also checked. For example, if a pixel is identified as skin tissue in the breast interior, it is rejected. The "clean up" procedure insures selection of tissue types that will most likely lead to a convergent solution.

V. NEURAL NETWORK CLASSIFICATION

The neural network classifier, proposed in Ref. 18, is also implemented with the CGEP method, and reconstructions are performed for the full-sized, 2D breast model. For details on neural network classification, see Ref. 18. Upon completion of the training state, the neural network configuration is 2–7–3, or 2 neurons in the first layer, 7 neurons in the second, and 3 neurons in the third. All activation functions are Log–Sigmoid. The initial learning rate is set to 0.09 and the momentum is set to 0.9. As with the BMC classifier, the NN classifier is used with the CGEP algorithm at every 10th iteration. A database is generated in the same manner as described in Ref. 18. The Matlab toolkit is used to train the neural network architecture.

VI. SIMULATION RESULTS

The CGEP, CGEP–BMC, and CGEP–NN methods were used to reconstruct the full-sized, 2D breast model discussed in Sec. II. Both noiseless and noisy cases were tested. By noiseless, we mean there is no signal noise in the scattered data and the mean values of the tissue parameters are used without variations. As in Ref. 9, for the noisy case, tissue losses are allowed to vary by 10% around the mean values, tissue sound speeds are allowed to vary by 2% (recall that sound speeds have considerably less impact on the imaginary part of the object than attenuation does), and the signal-to-noise ratio (SNR) of the scattered field measurements is set to 70 dB. The noise is meant to model both the system noise and point-to-point variations of the material parameters.

The three error measurements defined in Ref. 9 are the (1) relative residue error (RRE) which is defined as the normalized distance between the measured and computed scattered fields, (2) mean square error (MSE) which is defined as the normalized distance between the true and reconstructed objects, and (3) mean square error of the imaginary part (ImMSE) which is defined as the normalized distance between the imaginary part of the true and reconstructed objects. The latter is useful because only the imaginary part is used to distinguish a tumor from its surrounding tissue,

$$\mathrm{RRE}=\sqrt{\frac{\sum_{l=1}^{L}\|\hat{\vec{u}}_s-\vec{u}_s\|^2}{\sum_{l=1}^{L}\|\vec{u}_s\|^2}}, \tag{23}$$

$$\mathrm{MSE}=\frac{\|\hat{\vec{o}}-\vec{o}\|}{\|\vec{o}\|}, \tag{24}$$

$$\mathrm{ImMSE}=\frac{\|\mathrm{IMAG}(\hat{\vec{o}}-\vec{o})\|}{\|\mathrm{IMAG}(\vec{o})\|}, \tag{25}$$

where $\hat{o}$ indicates an estimated value.

The initial estimate used for the object function is the same one used in Ref. 9. The procedure for acquisition of the initial values is described in Ref. 9. Basically, they are obtained via two steps. The sound speed profile is acquired using either the Born approximation or the GN method for the real part of the object only—that is, the attenuation is set to zero. Once the sound speed profile has been calculated, the water, skin, fat, and glandular regions are located on the basis of pixel sound speeds. Termination of the algorithm occurs when $\mathrm{RRE}<1.0\times10^{-5}$ or when a maximum of 50 iterations has been performed. The material classifiers are introduced at the 10th iteration and are then used every 10 iterations thereafter. The timing for introduction of the classifiers is determined empirically.

All simulations were performed on a Cray T3E machine with 128 processors. Each simulation took from 3 to 6 wall-clock hours. A detailed flowchart for both algorithms is shown in Fig. 4.

A. Results without noise

Figure 5 shows the reconstructed images at iterations 10 and 50 for the CGEP method. At the 10th iteration, the tumors and cysts, even the minimal ones, are detectable. The reconstructed values at iteration 50 are quantitatively more accurate than those at iteration 10. Figure 6 shows the reconstructed results at iterations 10 and 40 for the CGEP–BMC

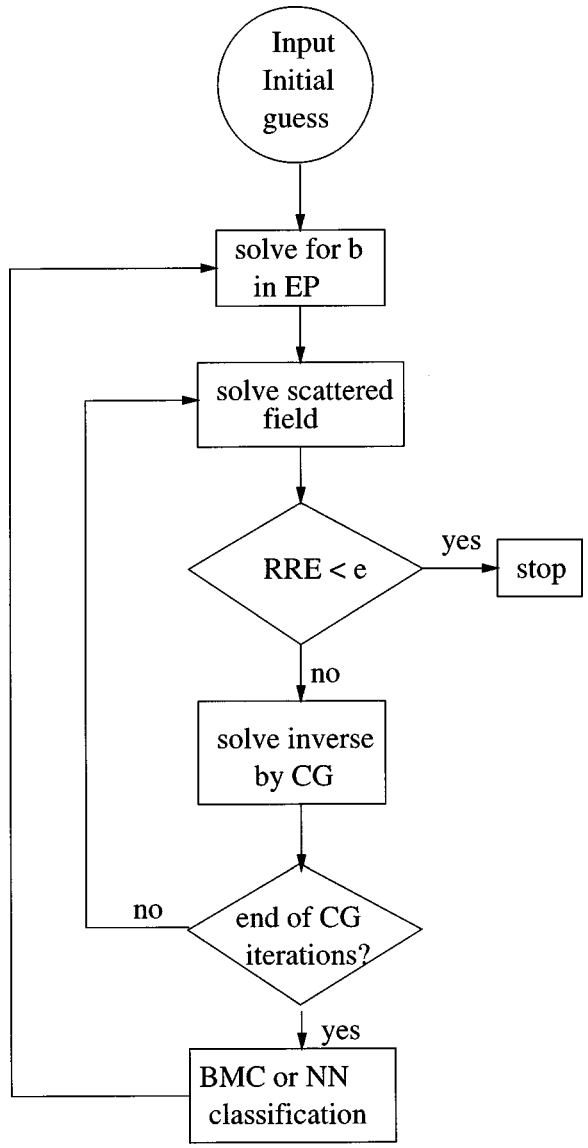

FIG. 4. Flowchart for the CGEP method with material classification.

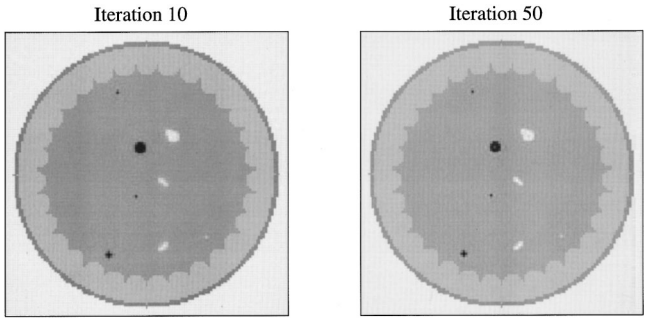

FIG. 5. Results fr the CGEP method without noise.

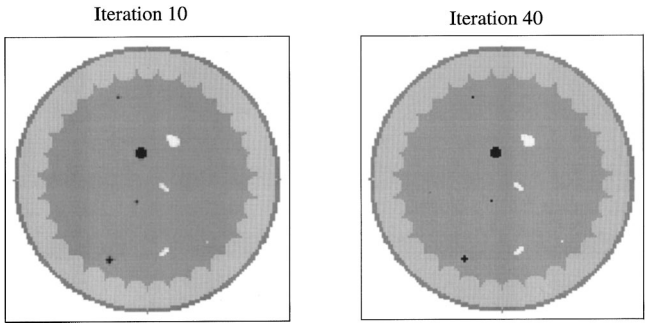

FIG. 6. Results for the CGEP–BMC method without noise.

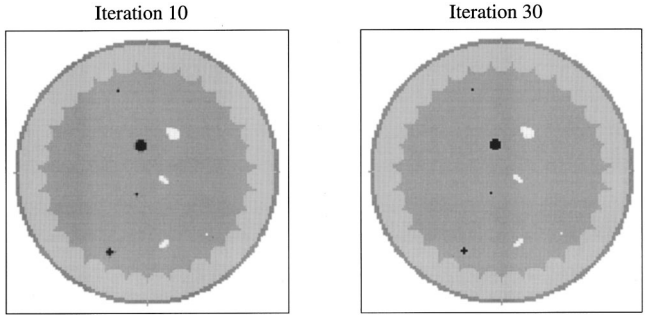

FIG. 7. Results for the CGEP–NN method without noise.

method. At the 10th iteration, some pixels are left unclassified, but all the unclassified pixels are correctly classified at the 40th iteration. Figure 7 shows the reconstructed images at the 10th and 30th iterations for the CGEP–NN method. At the 30th iteration, the entire image is correctly classified.

The RRE curves for the CGEP, CGEP–BMC, and CGEP–NN methods are shown in Fig. 8. For the CGEP method, the RRE values decrease monotonically. Application of the classifiers at intervals of 10 iterations causes the RRE values to drop substantially except at the 20th iteration. At the 20th iteration, classification actually causes an increase in the RRE values. The decrease of the RRE values is desirable while the increase after classification is not unexpected. During classification some pixels are misclassified, and this misclassification causes the increase. The same phenomenon occurs in the MSE and ImMSE curves shown in Fig. 9. As can be seen on the flowchart Fig. 4, the classified object is used as an initial guess for the next series of CGEP iterations. This restarting strategy is beneficial to the CGEP algorithm since it results overall in smaller errors. The convergence of the

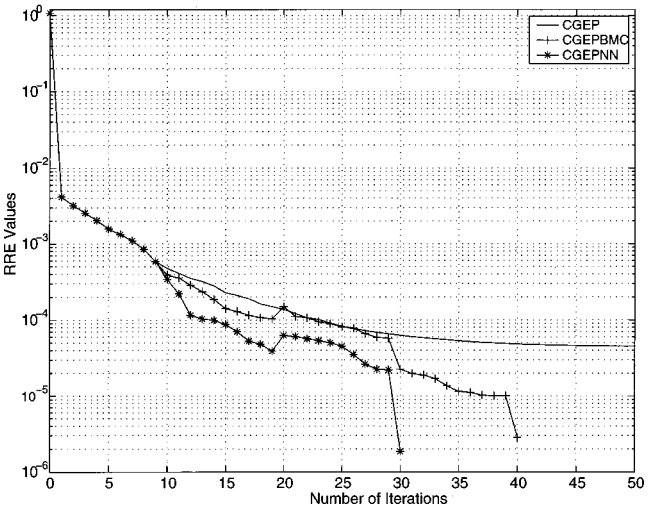

FIG. 8. RRE curves without noise.

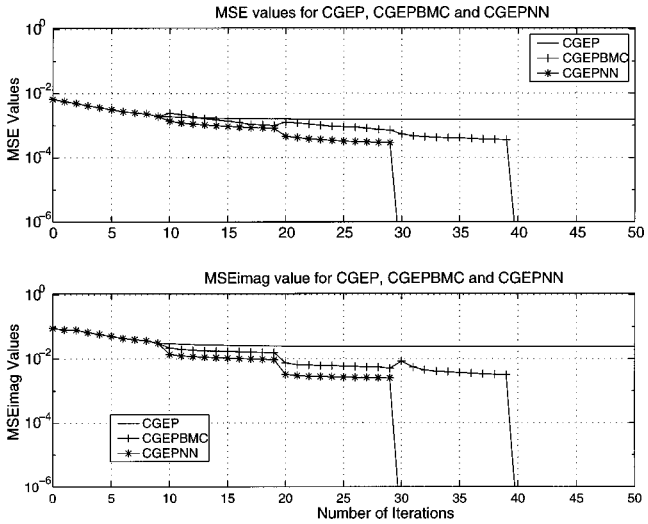

FIG. 9. MSE and ImMSE curves without noise.

CGEP method itself guarantees an eventual decrease in the error.

At the 40th iteration the RRE value for the CGEP–BMC method is 2.86×10^{-6}, and the RRE value for the CGEP–NN method at the 30th iteration is 1.85×10^{-6}. Both values are smaller than the predetermined criterion 10^{-5}. Thus the algorithms terminate.

Comparison of Figs. 5, 6, and 7 shows that the reconstructed images for the CGEP–BMC and CGEP–NN methods are more accurate than those of the CGEP method. Although both the CGEP–BMC and CGEP–NN methods correctly classify the object, the CGEP–NN method requires fewer iterations. For the noiseless case, the CGEP–BMC method requires 40% fewer iterations and the CGEP–NN method requires 55% fewer iterations than the CGEP method. Thus both the BMC and NN classifiers significantly reduce the number of iterations required and, hence, accelerate convergence.

Results for the CGEP–BMC and CGEP–NN methods are compared with those of the GN and GN–UMC methods reported in Ref. 9. Comparison of Figs. 6 and 7 with Fig. 4(b) in Ref. 9, which gives results for the GN–UMC method, shows that all three methods correctly classify the object, while the GN–UMC method takes the least number of iterations. In Table II we list the RRE, MSE, and ImMSE values after the final iteration of the CGEP, CGEP–BMC, CGEP–NN, GN, and GN–UMC methods. The RRE value for the CGEP method is an order of magnitude smaller than that of the GN method. Hence, reconstruction with the CGEP method is more accurate than with the GN method. In addition, the CGEP–BMC and CGEP–NN methods result in RRE values an order of magnitude smaller than that of the GN–UMC method.

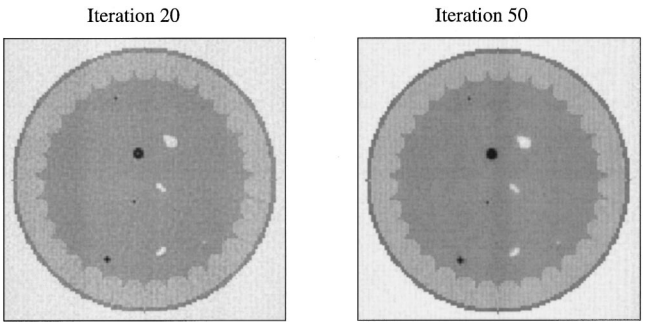

FIG. 10. Results for the CGEP method with noise.

B. Results with noise

Reconstruction results for the CGEP method with noise are shown in Fig. 10. Tumors and cysts are all detectable at both iterations 20 and 50. The reconstructed image at iteration 50 has smoother homogeneous regions than those at iteration 20. Figure 11 shows the reconstructed images at iterations 20 and 50 for the CGEP–BMC method. At iteration 20, notice that the minimal tumors and some pixels of the larger tumors are misclassified as glandular tissue, some cyst pixels are misclassified as glandular tissue, and some pixels are left unclassified. Some of the misclassified pixels are corrected at iteration 50. While all the tumors are correctly classified at iteration 50, the minimal cyst is still not. However, more pixels in the region of glandular tissue are left unclassified at iteration 20 than at iteration 50. Figure 12 shows the reconstruction images for the CGEP–NN method. Clearly all the tumors and cysts are correctly classified by both iterations 20 and 50. Also, there are more pixels left unclassified at iteration 20 than at iteration 50.

The RRE curves for the CGEP, CGEP–BMC, and CGEP–NN methods are shown in Fig. 13. Again we notice that the RRE values increase after classification has been applied. However, overall the RRE values decrease, although they decrease more slowly than those for the noiseless case. This is because the noisy case is more ill-posed. The CGEP–NN method results in a final value of RRE$=3.365\times10^{-4}$, and the CGEP–BMC results in RRE$=3.47\times10^{-4}$. Both are just slightly better than the final RRE value for the CGEP method RRE$=3.54\times10^{-4}$. The MSE and ImMSE curves for the CGEP, CGEP–BMC, and CGEP–NN methods are shown in Fig. 14. The values for the CGEP–NN method are smaller than for the CGEP and CGEP–BMC methods, indicating that reconstruction with the CGEP–NN is more accurate than with the CGEP–BMC method.

TABLE II. Error comparison for the CGEP, CGEP–BMC, CGEP–NN, GN, and GN–UMC methods without noise.

Error	CGEP	CGEP–BMC	CGEP–NN	GN	GN–UMC
RRE	4.527×10^{-5}	2.851×10^{-6}	1.858×10^{-6}	8.897×10^{-4}	1.914×10^{-5}
MSE	1.547×10^{-3}	0.0	0.0	1.726×10^{-3}	0.0
ImMSE	2.431×10^{-2}	0.0	0.0	1.726×10^{-2}	0.0

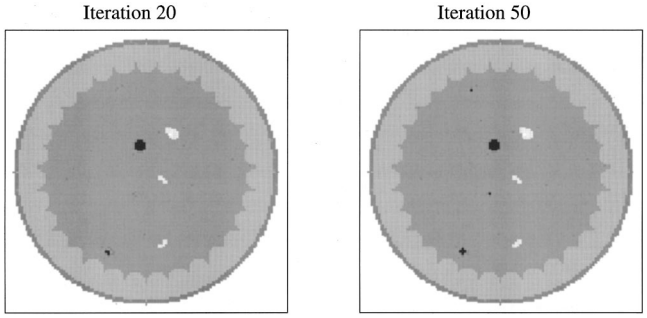

FIG. 11. Results for the CGEP–BMC method with noise.

Comparison of Figs. 11 and 12 with Fig. 7(b) in Ref. 9 shows that more pixels are misclassified with the GN–UMC method. The reconstructed images for the CGEP–BMC and CGEP–NN methods clearly have better visual contrast. The RRE, MSE, and ImMSE values for the last iteration of the CGEP, CGEP–BMC, and CGEP–NN methods and the best RRE, MSE, and ImMSE values for the GN and GN–UMC methods are listed in Table III. Although the RRE values for the CGEP–BMC and CGEP–NN methods are just slightly better than that of the GN–UMC method, the MSE values are an order of magnitude smaller than that of the GN–UMC method. This is also true for the ImMSE values. These results show that, as expected, the CGEP-based classification techniques are more robust to noise than the GN–UMC technique. With the CGEP method all error values decrease monotonically for both noiseless and noisy cases. However, when the GN method proposed in Ref. 9 is used for the noisy case, the ImMSE value at the ninth iteration is smaller than the value at the 20th iteration. The GN results diverge, while the CGEP results do not. The results shown are restricted to an SNR of 70 dB to facilitate comparison of the methods presented in this paper with the earlier GN–UMC method for which results are available only for 70 dB noise. In fact, it was found that the CGEP–NN method is able to detect tumors even at 40 dB.

C. Computational cost

The computational cost of the CGEP-based methods is dominated by the conjugate gradient iterations, which in turn are dominated by FFTs. The computational cost is $6N_{\mathrm{CGEP}}N_{\mathrm{fwd}}LN\log_2 N$, where N_{CGEP} is the product of the number of CG iterations and the number of edge preserving regularizations, N_{fwd} is the number of iterations for the forward solver, L is the number of sources, and N is the total number of cells into which the object is discretized. For a full-sized problem, N_{CGEP} is about 50 for an acceptable reconstruction and N_{fwd} is about 100. So the computational cost is $O(N^{5/2}\log_2 N)$. For the GN–UMC method, the computational cost is dominated by the GN method, whose computational cost as given in Ref. 8 is $4N_{\mathrm{GN}}(2N_{\mathrm{CG}}+N_{\mathrm{fwd}})LN\log_2 N$) where N_{GN} is the number of Gauss–Newton iterations, N_{CG} is the number of conjugate gradient iterations used to approximate the inverse of the Jacobian matrix, and the other parameters are the same as for the CGEP method. L, N_{GN}, and N_{CG} grow as the edge dimension of the array, i.e., $N^{1/2}$.[8] Thus the cost for the GN method is $O(N^{5/2}\log_2 N)$ which is the same as that of the CGEP method.

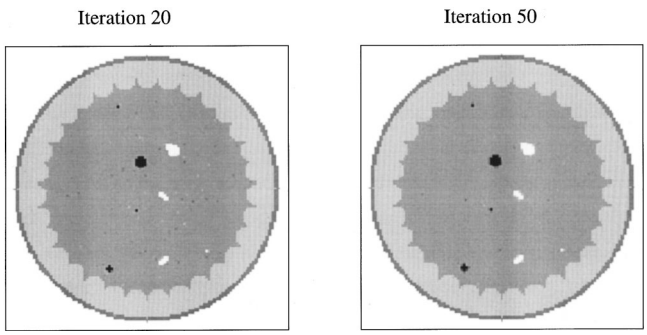

FIG. 12. Results for the CGEP–NN method with noise.

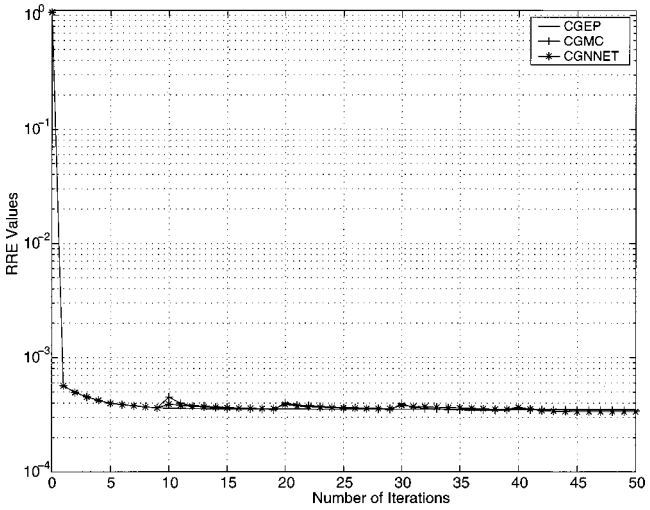

FIG. 13. RRE curves with noise.

VII. SUMMARY

In this paper, the CGEP method is used to solve the inverse problem for the detection of small breast tumors. The CGEP method is an iterative technique which is robust to noise but converges slowly. To accelerate convergence, a

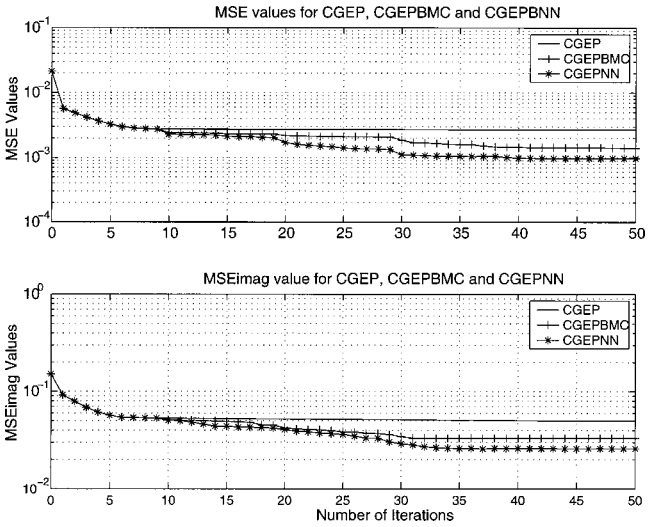

FIG. 14. MSE and ImMSE curves with noise.

TABLE III. Error comparison for the CGEP, CGEP–BMC, CGEP–NN, GN, and GN–UMC methods with noise.

Error	CGEP	CGEP–BMC	CGEP–NN	GN	GN–UMC
RRE	3.540×10^{-4}	3.470×10^{-4}	3.365×10^{-4}	3.593×10^{-4}	4.161×10^{-4}
MSE	2.702×10^{-3}	1.406×10^{-3}	9.707×10^{-4}	1.938×10^{-2}	1.920×10^{-2}
ImMSE	5.012×10^{-2}	3.302×10^{-2}	2.58×10^{-2}	0.1170	0.1008

new material classification technique is introduced as an extra step in the CGEP algorithm. The classifier, essentially a Bayesian classifier, is called bivariate material classification (BMC) since it simultaneously employs *a priori* knowledge of two breast tissue parameters, the sound speed and attenuation. The sound speed and attenuation for each type of tissue in the breast are assumed to be independently Gaussian distributed with known means but unknown variances. Histograms of the reconstructed images are used to estimate the variances and probability required by the classifier. A neural network (NN) classifier introduced in an earlier paper is used with the CGEP method as well. Both methods are used to reconstruct a full-sized, two-dimensional breast model for both noiseless and noisy cases. Numerical results show that the BMC and NN classifiers significantly reduce the number of iterations required for image reconstruction. In addition, the results are more accurate than that of the CGEP method without classification. For the noiseless case, the CGEP–NN method requires 55% fewer iterations and the CGEP–BMC method requires 40% fewer.

The CGEP-based methods are compared with each other as well as with the Gauss–Newton-based method (GN–UMC) proposed in Ref. 9. For the noiseless case, all three methods correctly classify the entire object. However, for the noisy case, the CGEP–BMC and CGEP–NN reconstructions are more accurate and yield smaller errors than those of the GN–UMC method. Thus the CGEP–BMC and the CGEP–NN methods are more robust to noise than the GN–UMC method. This is partly because information about both the sound speed and attenuation is used for classification in the CGEP–BMC and CGEP–NN methods, whereas only the attenuation is used for classification in the CN–UMC method. The CGEP-based methods do break down as the SNR is increased and start diverging at 60 dB, although the CGEP–NN method is still able to detect tumors even at 40 dB. The CGEP–BMC, CGEP–NN, and GN–UMC methods have comparable computational cost.

The CGEP–NN method gives slightly better results than the CGEP–BMC method. Differences between the two methods may be attributed to the fact that the neural network classifier is nonparametric, or distribution free, while the Bayesian classifier is parametric. For Bayesian classification to be correct, the accuracy of the distribution model is critical. However, the assumption that the sound speed and attenuation are independently Gaussian distributed is questionable. More time is required by the CGEP–BMC method than by the CGEP–NN method, although the CGEP–NN method requires a training state that can be time consuming. However, once the neural network architecture has been established, it can be re-used to perform classification.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation, Grant Nos. ECS-9711051 and EIA-9818212, and by the Carl M. Hansen Foundation.

[1] American Cancer Society, "Breast Cancer Facts and Figures—1997," Atlanta, 1-800-ACS-2345, 1997.
[2] Q. Zhu and B. D. Steinberg, "Wavefront amplitude distribution in the female breast," J. Acoust. Soc. Am. **96**, 1–9 (1994).
[3] R. J. Pauls and B. D. Steinberg, A refraction-based simulation to explore wavefront distortion in ultrasound mammography: Refraction-induced amplitude effects, Internal Report UP-14-92, Valley Forge Research Center, University of Pennsylvania, 1992.
[4] C. W. Manry and S. L. Broschat, "The FDTD method for ultrasound propagation through a two-dimensional model of the human breast," J. Acoust. Soc. Am. **94**, 1774–1785 (1993).
[5] W. C. Chew, Y. M. Wang, G. Otto, D. Lesselier, and J. C. Bolomey, "On the inverse source method of solving scattering problems," Inverse Probl. **10**, 547–553 (1994).
[6] W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," IEEE Trans. Med. Imaging **9**, 218–225 (1990).
[7] L. Liu, X. Zhang, and S. L. Broschat, "Ultrasound imaging using variations of the iterative Born technique," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**, 574–583 (1999).
[8] D. T. Borup, S. A. Johnson, W. W. Kim, and M.J. Berggren, "Nonperturbative diffraction tomography via Gaussian-Newton iteration applied to the scattering integral equation," Ultrason. Imaging **14**, 69–85 (1992).
[9] C. W. Manry and S. L. Broschat, "Inverse imaging of the breast with a material classification technique," J. Acoust. Soc. Am. **103**, 1538–1546 (1998).
[10] R. E. Kleinman and P. M. van den Berg, "A modified gradient method for two-dimensional problems in tomography," J. Comput. Appl. Math. **42**, 17–35 (1992).
[11] A. Roger, "Newton-Kantorovitch algorithm applied to the electromagnetic inverse problem," IEEE Trans. Antennas Propag. **29**, 232–238 (1981).
[12] A. Frenchois and C. Pichot, "Microwave imaging—complex permittivity reconstruction with a Levenberg-Marquardt method," IEEE Trans. Antennas Propag. **45**, 203–215 (1997).
[13] N. Joachimowicz, C. Pichot, and J.-P. Hugonin, "Inverse scattering: An iterative numerical method for electromagnetic imaging," IEEE Trans. Antennas Propag. **39**, 1742–1752 (1991).
[14] P. M. van den Berg, B. J. Kooij, and R. E. Kleinman, "Image reconstruction from Ipswich data—II," IEEE Antennas Propag. Mag. **39**, 29–32 (1997).
[15] P. Lobel, R. Kleinman, C. Pichot, L. Blanc-Feraud, and M. Barlaud, "Conjugate gradient method for solving inverse scattering with experimental data," IEEE Antennas Propag. Mag. **38**, 48–51 (1996).
[16] P. Lobel, L. Blanc-Feraud, C. Pichot, and M. Barlaud, "A new regularization scheme for inverse scattering," Inverse Probl. **13**, 403–410 (1997).
[17] C. Pichot, P. Lobel, L. Blanc-Feraud, M. Barlaud, K. Belkebir, J. M. Elissalt, and J. M. Geffrin, "Gradient and Newton Kantorovich methods for microwave tomography," in *Inverse Problems in Medical Imaging and Nondestructive Testing*, edited by H. W. Engl, A. Louis, and W. Rundell (Springer, New York, 1997), pp. 169–187.
[18] X. Zhang, S. L. Broschat, and P. J. Flynn, "A conjugate gradient-neutral network technique for ultrasound inverse imaging," to appear in J. Comput. Acoust.
[19] H. A. van der Vost, "BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems," SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **13**, 631–644 (1992).

[20] E. L. Madsen, J. A. Zagzebski, G. R. Frank, J. F. Greenleaf, and P.L. Carson, ''Anthropomorphic breast phantoms for assessing ultrasonic imaging performance and for training ultrasonographers: Part 2,'' J. Clin. Ultrasound **10**, 91–100 (1982).
[21] S. A. Johnson, F. Stegner, C. Wilcox, J. Ball, and M. J. Berggren, ''Wave equations and inverse solutions for soft tissue,'' Acoust. Imaging **11**, 409–424 (1982).
[22] A. Tikhonov, V. Arsenin, and F. John, *Solutions to Ill-Posed Problems* (Wiley, New York, 1997).
[23] R. F. Harrington, *Field Computation by Moment Methods* (MacMillan, New York, 1968).
[24] D. T. Borup and O. P. Gandhi, ''Calculation of high-resolution SAR distributions in biological bodies using the FFT algorithm and conjugate gradient method,'' IEEE Trans. Microwave Theory Tech. **33**, 417–419 (1985).
[25] X. Zhang, S. L. Broschat, and P. J. Flynn (unpublished).
[26] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, ''Deterministic edge-preserving regularization in computed imaging,'' IEEE Trans. Image Process. **6**, 298–311 (1997).
[27] D. Geman and G. Reynolds, ''Constrained restoration and the recovery of discontinuity,'' IEEE Trans. Pattern Anal. Mach. Intell. **14**, 367–383 (1992).
[28] R. Gonzales and R. Woods, *Digital Image Processing* (Addison-Wesley, New York, 1992).