Bayesian compressive sensing for ultrawideband inverse scattering in random media

Ahmed E Fouda\textsuperscript{1} and Fernando L Teixeira\textsuperscript{2}

\textsuperscript{1}Halliburton Energy Services, Sensor Physics and Technology, Houston, TX, USA
\textsuperscript{2}ElectroScience Laboratory, Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, USA

E-mail: fouda.1@osu.edu and teixeira@ece.osu.edu

Received 23 March 2014
Accepted for publication 11 June 2014
Published 28 October 2014

Abstract
We develop an ultrawideband inverse scattering technique for reconstructing continuous random media on the basis of Bayesian compressive sensing. In addition to providing maximum \textit{a posteriori} estimates of the unknown weights, Bayesian inversion provides an estimate of the confidence level of the solution, as well as a systematic approach for optimizing subsequent measurement(s) to maximize information gain. We impose sparsity priors directly on spatial harmonics to exploit the spatial correlation exhibited by continuous media and then efficiently solve for their posterior probability density functions by using a fast relevance vector machine. We linearize the problem by using the first-order Born approximation, which enables us to combine, in a single inversion, measurements from multiple transmitters and ultrawideband frequencies. To extend the method to high-contrast media, we also introduce a Bayesian compressive sensing distorted-Born iterative approach. Further, we illustrate the use of time-reversal strategies to adaptively focus the inversion effort onto subdomains of interest and hence reduce the overall inversion cost. The proposed techniques are applied to a number of scenarios including crosshole and borehole sensing.

Keywords: Bayesian, compressive sensing, ultrawideband

(Some figures may appear in colour only in the online journal)
The goal of inverse scattering is to estimate unknown parameters of target(s) of interest from noisy (cluttered) measurements. Target parameters may include location, size, orientation, and material properties [1–6]. Electromagnetic inverse scattering finds many applications in medical imaging [7–12], through-wall imaging [13–15], non-destructive testing [16–18], ground-penetrating radar [19–22], and geophysical exploration in general [3, 4, 23, 24].

In Bayesian-based inversion, both target parameters and clutter are modeled as random variables with certain probability density functions (PDFs). The inversion algorithm combines (any) a priori information on the target’s parameters with physics-based forward-problem PDFs and array acquisitions to produce a posterior PDFs of the unknowns [1, 2, 9, 25–29]. This approach provides a means for measuring the confidence interval of the inversion and adaptively optimizing subsequent measurement(s) [30, 31]. A Bayesian inference applied to compressive sensing was presented in [31, 32], where sparsity priors were imposed on a compressible (sparse) set of unknowns. That problem was solved efficiently by using the relevance vector machine (RVM) technique [33, 34].

Recently, Bayesian compressive sensing (BCS) has been applied in microwave imaging of sparse discrete scatterers by using single frequency data and the contrast source formulation [28] or the first-order Born approximation [29]. Bayesian compressive sensing, combined with signal subspace methods, for imaging discrete targets has been presented in [35–37]. Bayesian compressive sensing was recently also used in [38–40] for mapping obstacles with see-through capability in a multimodal sensing scenario.

In non-Bayesian inverse scattering techniques, a cost function, to be iteratively minimized, is defined, and optimization techniques such as the conjugate gradient method [41] are used to guide the iterations [3, 4, 42–45]. Bayesian inversion alleviates the need for that since it has a ‘built-in’ measure for accuracy through the confidence level it provides. In addition, BCS solved by using the RVM, as presented in [31, 33], provides an elegant, closed-form solution for the posterior PDF; therefore, (costly) numerical computations of higher-order integrations otherwise done by using Markov Chain Monte Carlo and Gibbs sampling [46, 47], as in [1, 2, 9, 24], for example, are unnecessary. Moreover, BCS with RVM does not require inversion of the projection matrix relating measurements to model parameters. Note that this matrix may not be square where, for example, the number of measurements is less than the number of unknowns, and it can be ill-conditioned, making it highly sensitive to noise.

In this paper, we develop BCS-assisted ultrawideband (UWB) inverse scattering techniques that exploit frequency decorrelation of the UWB interrogating signals to produce a statistically stable inversion, which does not depend on the particular realization of the (random) clutter/noise but only on its statistical properties [48, 49]. The new contributions of this paper can be summarized as follows: (i) We apply a linear regression model [28, 31, 33] to the electromagnetic inverse scattering problem under a first-order Born approximation, as in [29], but we extend it to incorporate, in a single inversion, UWB multistatic measurements, and we apply it to continuous random media rather than to (sparse) discrete scatterers. We present several examples based on numerical simulations to show the capability of the proposed method not only in reconstructing continuous random media properties, but also in estimating the confidence level of the inversion. (ii) We present an adaptive sensing approach for determining successive measurement locations so as to maximize the differential information gain from each measurement. (iii) We introduce a new technique, denoted time-reversal-assisted localized-inversion (TRALI), to reduce the computational cost of the inversion algorithm through adaptive focusing of the inversion effort onto sub-domains of
interest. (iv) We extend BCS inversion to high-contrast media (i.e., those that do not conform with the first-order Born approximation) by introducing a Bayesian compressive sensing distorted-Born iterative method (BDBIM). We compare the performance of BDBIM with that of conventional (non-Bayesian) DBIM and show the superior performance of the former. (v) Finally, we apply the proposed BCS inversion to layered and structured media (i.e., those comprising piecewise-constant contrast objects) to show that our method is not limited to smoothly varying continuous media.

It should be pointed out that the term ‘compressive sensing’ is used here in a broad sense to refer to problems where the unknown function (target locations and properties in our case) can be expressed as a sparse set of weights w.r.t. some expansion bases. This is done to conform to the prior usage in [28, 29] and does not match a more formal usage of the term that refers to recovering certain signals from sparse data acquisitions (i.e., using fewer samples or measurements than dictated by the Nyquist criterion) [50, 51]. In particular, in this work as well as in [28, 29], the number of measurements can be larger than the number of unknowns.

2. Bayesian compressive sensing using the relevance vector machine

Consider a linear regression model, where a vector \( \mathbf{y} \) of \( N \) noisy measurements is related to a vector \( \mathbf{w} \) of \( M \) (unknown) weights through the linear relationship

\[
\mathbf{y} = \mathbf{Bw} + \mathbf{n}
\]

where \( \mathbf{B} \) is the projection matrix and \( \mathbf{n} \) is a vector of additive noise. We are seeking maximum a posteriori (MAP) estimates for the weights as follows

\[
\hat{\mathbf{w}} = \arg \max_{\mathbf{w}} p(\mathbf{w} \mid \mathbf{y})
\]

The main challenge is trying to avoid ‘over-fitting’ the noisy measurements [33]. From Bayes’ rule, the posterior PDF is given by

\[
p(\mathbf{w} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{w})p(\mathbf{w})}{p(\mathbf{y})}
\]

Let us consider each term of the preceding PDF; we start with the prior \( p(\mathbf{w}) \). If we have a priori knowledge that the weights vector is sparse, meaning that only a few weights are non-zero, then a reasonable choice for the prior will be a sparsity prior such as the Laplace PDF. However, using such a prior, a closed-form solution for the posterior PDF cannot be obtained [31]. An alternative approach, introduced in [33], is to use hierarchical priors by defining \( p(\mathbf{w}) \) through a vector of hyperparameters \( \alpha \) as follows

\[
p(\mathbf{w}) = \int p(\mathbf{w} \mid \alpha)p(\alpha) \, d\alpha
\]

where the conditional PDF is defined as

\[
p(\mathbf{w} \mid \alpha) = \prod_{i=1}^{M} N(w_i \mid 0, \alpha_i^{-1})
\]

in which the hyperparameters are the reciprocals of the variances of the zero-mean normal distributions. The hyperparameters themselves are assumed to be distributed according to the following gamma distribution
with \(a\) and \(b\) being the scale parameters of the gamma distribution. The resulting prior in (4) is a student-\(t\) distribution that, with appropriate choice of scale parameters, is highly peaked at zero, thus favoring sparsity [33]. Now, consider the likelihood \(p(y|w)\). Assuming independent, zero-mean, Gaussian noise with variance \(\sigma_n^2\), the likelihood can be written as

\[
p(y|w) = \int p(y|w, \sigma_n^2)p(\sigma_n^2)d\sigma_n^2
\]

where

\[
p(y|w, \sigma_n^2) = \left(2\pi\sigma_n^2\right)^{-N/2}\exp\left(-\frac{1}{2\sigma_n^2} \|y - Bw\|^2\right)
\]

and the reciprocal of the noise variance is distributed according to the following gamma distribution with parameters \(c\) and \(d\)

\[
p(\sigma_n^2) = \Gamma(\sigma_n^2 | c, d)
\]

Combining (4), (7) and (8), the posterior becomes

\[
p(w | y) = \int\int p(y|w, \sigma_n^2)p(\sigma_n^2)p(w | \alpha)p(\alpha)d\sigma_n^2 d\alpha
\]

which can be simplified to

\[
p(w | y) = \int\int p(w, \alpha, \sigma_n^2 | y)d\alpha d\sigma_n^2
\]

\(p(w, \alpha, \sigma_n^2 | y)\) is the joint posterior PDF of all unknowns and can be factorized as follows

\[
p(w, \alpha, \sigma_n^2 | y) = p(w | y, \alpha, \sigma_n^2)p(\alpha, \sigma_n^2 | y)
\]

The first term in the rhs of (12) can be expanded as

\[
p(w | y, \alpha, \sigma_n^2) = \frac{p(y|w, \sigma_n^2)p(w | \alpha)}{p(y | \alpha, \sigma_n^2)}
\]

where

\[
p(y | \alpha, \sigma_n^2) = \int p(y|w, \sigma_n^2)p(w | \alpha)dw = \exp\left[-\frac{1}{2}(y^T C^{-1}y)\right]
\]

with \(C := \sigma_n^2 I + B\Sigma^{-1}B^T\), and \(A := \text{diag}(\alpha)\). Using (8), (5) and (14) in (13),

\[
p(w | y, \alpha, \sigma_n^2) = (2\pi)^{-\left(M+1)/2\right}} |\Sigma|^{-1/2} \times \exp\left(-\frac{1}{2}(w - \mu)^T \Sigma^{-1}(w - \mu)\right)
\]

where \(\Sigma = (\sigma_n^{-2}B^T B + A)^{-1}\) and \(\mu = \sigma_n^{-2}\Sigma B^Ty\). This is the sought posterior PDF of the weights once the hyperparameters \(\alpha\), which are embedded in \(A\), and the noise variance \(\sigma_n^2\) are determined. Towards this end, the second term in the rhs of (12) can be approximated by an impulse centered around the most probable (MP) values of \(\alpha\) and \(\sigma_n^2\), as follows

\(^3\) A reasonable choice, adopted in [33], is to set scale parameters \(a\) and \(b\) to zero. In this case, \(p(\ln(\alpha))\) is uniform, i.e. the hyperparameters become scale invariant.
Note that \( p(\mathbf{a}, \sigma^2_n | y) \propto p(y | \mathbf{a}, \sigma^2_n) p(\mathbf{a}) p(\sigma^2_n) \), and by properly adjusting the scale parameters \( a, b, c \) and \( d \), \( \{|a|_{\text{MP}}\} \) and \( \{|\sigma^2_n|_{\text{MP}}\} \) can be obtained by maximizing only the marginal likelihood \( p(y | \mathbf{a}, \sigma^2_n) \), i.e.

\[
\left( \{|a|_{\text{MP}}\}, \{|\sigma^2_n|_{\text{MP}}\} \right) = \arg \max_{a, \sigma^2_n} \left\{ \frac{\exp\left( -\frac{1}{2} (y^T \mathbf{C}^{-1} y) \right)}{(2\pi)^{N/2} |\mathbf{C}|^{1/2}} \right\}
\]

This is known as type-II maximum-likelihood process and can be efficiently solved by using the fast relevance vector machine (RVM) presented in [34]. Finally, from (15), the MAP estimates and the corresponding covariance matrix are given by

\[
\hat{\mathbf{w}} = \mathbf{\mu} \left( \{|a|_{\text{MAP}}\}, \{|\sigma^2_n|_{\text{MAP}}\} \right)
\]

and

\[
\text{cov}(\mathbf{w}) = \mathbf{\Sigma} \left( \{|a|_{\text{MAP}}\}, \{|\sigma^2_n|_{\text{MAP}}\} \right)
\]

### 3. UWB inverse scattering based on the Born approximation

So far, we have discussed a BCS solution for a generic linear regression model relating noisy measurements to sparse weights. To apply this model to the electromagnetic (EM) inverse scattering problem, we first have to be able to write the scattered field as a linear combination of the model weights. One way to do this is by using the contrast source formulation presented in [28]. Another way is by using the first order Born approximation as presented in [29]. In this work, we choose the latter approach for two reasons that will be clarified shortly.

Under the first-order Born approximation and assuming two-dimensional problem for simplicity, the scattered field at spatial location \( \mathbf{r} \) and frequency \( \omega_k \) resulting from incident field \( E_{t}^{\text{inc}} \) generated by transmitter \( t \) is given by

\[
E_{s}^{\text{inc}}(\mathbf{r}, \omega_k) = \int_{D} \tau(\mathbf{r}', \omega_k) E_{t}^{\text{inc}}(\mathbf{r}', \omega_k) G(\mathbf{r}, \mathbf{r}', \omega_k) d\mathbf{r}'
\]

where \( G \) is the 2D scalar Green’s function and \( D \) is the support of the scattering object. \( \tau \) is the complex contrast function given by

\[
\tau(\mathbf{r}, \omega_k) = \left[ \varepsilon_{r}(\mathbf{r}) - \langle \varepsilon_{r} \rangle \right] - j \left[ \frac{\sigma(\mathbf{r}) - \langle \sigma \rangle}{\omega_k \varepsilon_0} \right] = \Delta \varepsilon(\mathbf{r}) - j \frac{1}{\omega_k \varepsilon_0} \Delta \sigma(\mathbf{r})
\]

with \( \varepsilon \) and \( \sigma \) being the relative permittivity and conductivity, respectively, and \( \langle \varepsilon \rangle \) and \( \langle \sigma \rangle \) the corresponding mean values of the background medium. Real and imaginary parts of the scattered field, recorded at \( N_s \) sensors, can be stacked in a column vector as follows.
where

\[ g_{r,\omega_k}(r, \omega_k) = \int_{D_p} E^{inc}_t(r', \omega_k) G(r, r', \omega_k) dr' \tag{24} \]

Equation (20) can now be written in matrix form as follows

\[ e_{t, k} = G_{t,k} (F^{-1}t) + n_{t,k} \tag{25} \]

where \( F^{-1}t \) is a real-valued contrast vector, given by

\[ F^{-1}t = \begin{bmatrix}
\Delta\varepsilon_t(r_1) \\
\vdots \\
\Delta\varepsilon_t(r_{N_p}) \\
\Delta\sigma(r_1) \\
\vdots \\
\Delta\sigma(r_{N_p})
\end{bmatrix} \tag{26} \]

\( F^{-1} \) is the inverse Fourier transform matrix, and \( t \) is the vector of the spatial harmonics of the real contrast function. \( G_{t,k}F^{-1} \) can now be perceived as the projection matrix \( B \) in (1) and the unknown weights are the spatial harmonics \( t \). Once the covariance matrix of \( t \) is solved for by using the RVM, the covariance matrix of the real contrast vector can be computed as \( \text{cov}(F^{-1}t) = F^{-1}\text{cov}(t)(F^{-1}t)^T \).

Formulating the problem in terms of the spatial harmonics, rather than the contrast function itself, has the following advantages: 1. Spatial harmonics conform better with the sparsity requirement of the model, since the solution is likely to possess an amount of spatial correlation that would make the contrast function more sparse in the spatial harmonics domain than in the spatial domain. 2. Spatial harmonics provide better regularization for the
solution; for example, one can choose to solve for a subset of the spatial harmonics according to the problem specifics and/or the available resources.

Measurements corresponding to illuminations from different transmitters can be stacked in one inversion and, assuming negligible dispersion over the frequency band used, multi-frequency measurements can be stacked in the same inversion as well (since $t$ is frequency independent under this assumption). This yields what we call multistatic UWB Bayesian inversion. In UWB inversion, low frequencies are more sensitive to lower spatial harmonics, whereas high frequencies are more sensitive to higher spatial harmonics. The highest spatial harmonic that can be resolved depends on the maximum frequency that can be used without violating the Born approximation. Increasing the number of uncorrelated measurements, whether from sufficiently spaced sensors and/or frequency samples, makes the inversion statistically more stable against random noise and/or clutter [48, 49].

The ability to define transmitter- and frequency-independent unknowns is a consequence of the adopted Born approximation. In the contrast source formulation in [28], the unknowns are the equivalent currents of the contrast function, which vary with frequency and incident field. For high-contrast media, iterative inversion approaches can be used, as discussed in section 6 ahead.

4. Results

The UWB BCS inversion process is summarized in figure 1. The pointwise distribution of the medium can be characterized by certain parameters such as the average electrical properties, the correlation length, and the contrast level of the fluctuations. Another parameter dictated by the problem is the signal-to-noise ratio (SNR) of the measurements. A priori knowledge of any of the medium parameters can be used to select the user-controlled parameters. Those include the domain of interest (DOI) pixelization, the frequency band used, and the number and location of sensors. The UWB BCS inversion is then invoked, and the output contrast level and confidence level are used to decide whether more iterations are needed. In that case, posterior information can be fed back to refine the user-controlled parameters of the next iteration.

An example of the permittivity contrast of a continuous random medium is shown in figure 2(a). This distribution is a realization of a Gaussian random process with zero-mean, standard deviation of 0.064, and a Gaussian correlation function with correlation length
The spatial spectrum is shown in figure 2(b). For simplicity, we consider 2D models throughout this work, but evidently the same analysis can be easily extended to 3D cases. The background medium is assumed to have mean permittivity of 3 and mean conductivity of 0.15 mS/m. This example may correspond to underground imaging of dry soil [3]. The interrogating frequency band ranges from 5 to 250 MHz with 50 samples. Note that the maximum wavenumber of the interrogating signal $k_{\text{max}} = 1.48 \times 2\pi$ rad/m is larger than the maximum spatial harmonic of the medium ($=1.1\times 2\pi$ rad/m computed across the diagonal). Forward problem simulations are carried out by using the finite-difference time-domain method [52].

In the inverse problem, the DOI is discretized uniformly into 20 × 20 pixels, and the 2D Green’s function is computed analytically by assuming known average medium properties. We use $N_s = 15$ multistatic sensors deployed either in full-aspect (FA) circular geometry as shown in figure 3(a), crosshole (CH) geometry as in figure 3(d), or borehole (BH) geometry as in figure 3(g). Transmitters are point sources in 2D (infinite line source) radiating TM$_z$ polarization. For the particular application of underground imaging, the $x−y$ plane in the FA case can be perceived as the horizontal plane, with the distribution shown being a horizontal cross-section in the formation and the sensors deployed in circularly distributed wells. For the BH and CH cases, the distribution shown is a vertical cross-section, and the sensors are deployed in one or two wells, respectively. The SNR is assumed to be 10 dB for all measurements performed by using different sensors and frequencies.

Reconstructed profiles for the three geometries are shown in figures 3(b), (e), and (h), and the estimated standard deviations (which determine the confidence level of the inversions) are shown in figures 3(c), (f), and (i), respectively. Reconstructed images are interpolated to a finer grid for the sake of visualization. Comparing actual and reconstructed profiles, we note that FA and CH outperform BH. Also, the estimated standard deviation provides a reasonably good measure for the inversion accuracy, more obviously in the BH case, where the reduced-accuracy inversion in the right half of the investigation domain (farther from the array) is associated with higher standard deviation. Roughly speaking, the inversion accuracy and the reciprocal the of standard deviation at a certain point are proportional to the spatial resolution offered by the sensor array at that point.
4.1. Performance analysis

In this section, we provide some qualitative measures for assessing inversion accuracy and efficiency. We first define the actual rms error as

$$\text{Actual rms error} = \sqrt{\text{avg}_D \left| \hat{\epsilon}_r(r) - \epsilon_r(r) \right|^2}$$  \hspace{1cm} (27)

where $\hat{\epsilon}_r$ is the estimated permittivity. The (average) estimated standard deviation can be defined as

$$\text{Estimated std. deviation} = \sqrt{\text{avg}_D \left[ \text{diag} \left( \text{cov} \left( F^{-1} t \right) \right) \right]}.$$  \hspace{1cm} (28)

Both actual and estimated errors are plotted in figure 4(a) for the three previously discussed geometries and two SNRs. Percentile error is defined as the ratio of the absolute error to the rms of the actual contrast function ($\sqrt{\text{avg}_D \left| \epsilon_r(r) \right|^2}$). This plot shows that estimated error follows the actual error quite well, with the latter being always larger. This makes perfect
sense, since the estimated error accounts only for errors due to additive noise, whereas actual error includes in addition to noise, errors due to the adopted Born approximation and discretization error. Estimated SNRs are shown in figure 4(b). They are below their actual values by 1–2 dB, which indicates that the RVM solver overestimated the noise variance.

Errors and processing times for a FA array with uniformly distributed increasing number of sensors are tabulated in table 1. As expected, errors decrease monotonically as the number of sensors increases at the expense of increasing the processing time. Listed times are those required for solving the fast RVM using non-optimized Matlab code, running on a machine with average CPU speed of 2.7 G.cycle/s. They are very short times (almost real-times) w.r.t. the size and the number of measurements of the considered problem. Note that there are costs associated with computing the Green’s function and constructing the projection matrix, but those are considered as pre-processing costs.

Note also that in the examples shown, the only source of clutter is the additive white Gaussian noise (the disordered medium itself is not considered clutter since it is the medium to be reconstructed). To verify statistical stability, inversion was repeated for several realizations of the random noise and results very close to those in figures 3 and 4 were obtained. In scenarios where additional clutter arises from spurious scatterers (e.g., parts of the medium outside the domain of interest), frequency decorrelation will not be perfect and will be dependent on the medium properties. Nevertheless, using a large number of spatially spaced measurements and a sufficiently wide bandwidth will make the inversion less dependent on the particular realization of the clutter; and rather, the inversion will depend only on its statistical properties [48, 49].

Another measure for quantifying the confidence level of the inversion is the differential entropy (DE) [53]. Referring to (1), the DE of the posterior multivariate Gaussian PDF is given by

\[
h(p) = -\int p(w \mid y) \ln(p(w \mid y)) \, dw = \frac{1}{2} \ln \left( \frac{1}{(2\pi)^{M/2}} |\Sigma| \right)
\]

The DE given by the preceding equation is in information units (nats). It can be divided by \ln (2) to give the DE in bits. Recall that DE indicates the degree of ‘spreadness’ of a PDF, that is, random variables with PDF concentrated on a small interval yield smaller DE. For continuous random variables, DE can be negative (as opposed to the entropy of discrete random variables, which is always positive). Differential entropies for the setups of figure 4(a) are summarized in table 2. Individual values of DE do not give much information about the randomness of the PDF; however, comparing DEs of two setups gives an idea about the accuracy gained or lost (measured in units of information) on going from one setup to the other. This measure agrees well with the behavior described in figure 4(a).

4.2. Adaptive sensing

Our goal in this section is to develop a systematic procedure for optimizing the location(s) of subsequent measurement(s) such that the information gain from each measurement is maximized [30, 31]. The DE, after adding the \((N + 1)\)th measurement, can be written in terms of the DE of \(N\) measurements as follows [31]

\[
h(r_{\text{new}}) = h(p) - \frac{1}{2} \ln \left( 1 + \sigma_{T}^{-2} r_{B,N+1}^{T} \Sigma r_{B,N+1} \right)
\]

where \(r_{B,N+1}^{T}\) is the new row added to the projection matrix \(B\) associated with the \((N + 1)\)th measurement. To maximize information gain, the absolute value of the second term in the rhs of (30) should be maximized, which implies that \(r_{B,N+1}^{T}\) should be chosen such that
is maximized. In other words, we choose to place the next sensor where we expect the highest uncertainty in the measurement, in this way, the information gain is maximized [31]. The preceding equation is maximized by choosing $r_{b,N+1}$ to be the eigenvector of $\Sigma$ corresponding to the largest eigenvalue [31].

Referring to our case study, we apply the adaptive scheme to place new sensors in a ‘myopic’ sense (one sensor in each step) as shown in figure 5. Suppose that the locations of five sensors in step 1 are pre-determined; the goal then is to optimally place five more sensors. Also, suppose that sensors can be deployed only on a circle with 7.5 m radius. The figure shows the reconstructed profile from each step, the location of the sensors used, the estimated standard deviation, and the optimized projection vector. The new sensor has to be placed such that the field pattern produced from it best matches the optimized projection vector. Note that
Figure 5. Reconstructed profiles, standard deviations, and projection vectors for six steps of an adaptive sensing scenario. The letter ‘x’ indicates sensor locations in each step. SNR = 10 dB.
the standard deviation distribution is not enough to determine the location of the new sensor without computing its eigenvalue decomposition. Actual and estimated errors as well as the DE of each step are summarized in Table 3.

For comparison, a non-adaptive scenario is shown in Figure 6, where the same five sensors are pre-determined and the other five sensors are uniformly distributed as shown. Corresponding performance parameters are shown in Table 3 as well. From this comparison, it is obvious how adaptive optimized sensing yields more accurate inversion given the same number of sensors, or in other words, adaptive sensing can achieve a given inversion accuracy with fewer number of sensors.

5. Time-reversal-assisted localized inversion

If our interest is to reconstruct only a localized region of the investigation domain that can change dynamically, time-reversal (TR) focusing [54–57] can be used to achieve accurate localized inversion with a significantly shorter processing time. We call this technique ‘Time-Reversal-Assisted Localized-Inversion’ (TRALI). In TRALI, measurements from different sensors are linearly combined as follows

| Step | Actual error | Estimated error | Differential entropy (kb) |
|------|--------------|----------------|--------------------------|
| 1    | 0.117        | 0.108          | –2.32                    |
| 2    | 0.067        | 0.042          | –2.56                    |
| 3    | 0.041        | 0.025          | –2.82                    |
| 4    | 0.028        | 0.015          | –3.03                    |
| 5    | 0.023        | 0.012          | –3.12                    |
| 6    | 0.021        | 0.01           | –3.18                    |
| Non-adaptive sensing | 0.031 | 0.014 | –3.01 |

Figure 6. Non-adaptive sensing scenario. N_s = 10 and SNR = 10 dB.
\[ E^{TR}(r_p, \omega_k) = \sum_{i=1}^{N_k} \sum_{j=1}^{N} G(r_p, r_r, \omega_k) \left[ G^\dagger(r_r, r_p, \omega_k) E_i^s(r_r, \omega_k) \right] \]

(32)

where \( G^\dagger(r_r, r_p, \omega_k) \) is the complex conjugated Green’s function between pixel \( p \), in the region of interest, and location \( r \). Note that complex conjugation in the frequency domain is equivalent to TR. Assuming multistatic acquisition, the preceding equation is equivalent to simultaneously firing all transmitters to illuminate the DOI by a beam focused at location \( r_p \). Backscattering is then recorded by all receivers, time-reversed, and projected on pixel \( p \). In this way, \( E^{TR}(r_p, \omega_k) \) will be most sensitive to the contrast of pixel \( p \); consequently, using \( E^{TR}(r_p, \omega_k) \) in place of \( E_i^s(r_r, \omega_k) \) in the linear regression model (25) yields an accurate localized inversion. Of course, the rows of the projection matrix need to undergo the same linear combination in (32). An example is shown in figure 7. A thirty-transceiver FA array is used to obtain very accurate inversion of the entire DOI as shown in figures 7(a) and (b). Using the same array, TRALI is applied to obtain localized inversion of one hundred pixels in the upper right quarter of the DOI, as shown in figures 7(c) and (d). Note that the local inversion sub-domain does not need to be static or contiguous; also, it can be extended to encompass the entire DOI. To further assess the performance of TRALI, a ten-transceivers FA array (which has the same number of data points and requires almost the same processing time as TRALI) is used in figure 7(e) and (f). Corresponding total error, local error (of the upper right quarter), and processing time are summarized in table 4. TRALI is shown to produce local inversion with almost the same accuracy as the full multistatic acquisition, but with much less processing time. This comes at the expense of sacrificing accuracy elsewhere outside the local domain of interest. Using the same number of multistatic acquisitions as the TR focusing pixels results in a larger local error, but less overall error.
6. Bayesian distorted-Born iterative method

So far, we have considered the application of the proposed UWB BCS inversion to low-contrast media obeying the first-order Born approximation. In this section, we extend the applicability of the method to high-contrast continuous media. The proposed Bayesian compressive sensing inversion scheme can be applied iteratively, yielding what we call the ‘Bayesian compressive sensing distorted-Born iterative method’ (BDBIM). In conventional DBIM [42–45, 58, 59], a cost function is defined, usually as the $L_2$ norm between measured scattered field and synthetic scattered field computed from the reconstructed profile, and the method proceeds iteratively to minimize that cost function. A reconstructed profile from each iteration is used to compute the synthetic scattered field as well as the Green’s function used in the next iteration. The method converges when the cost function gets below a certain predetermined threshold. Although BDBIM proceeds the same way, instead of explicitly defining a cost function on the scattered fields, the estimated standard deviation from the Bayesian solver can be used as a stopping criterion. To illustrate that, consider the example in figure 8. In the first iteration, a uniform homogeneous background is used in the BCS inversion. The reconstructed profile is shown in figure 8(b). Actual error and estimated standard deviations are shown in figures 9(a) and (b), respectively.

The percentile error shown in figures 9(b) is the ratio of the estimated standard deviation to the rms of the contrast function contributed from each iteration. The reconstructed profile from the first iteration is plugged into a forward problem numerical solver and used to compute the synthetic scattered field and the Green’s function to be used in the following iteration. The synthetic scattered field is subtracted from the (noisy) measurements, and that differential signal is used as the measurements vector in the second iteration. The reconstructed profile from the second iteration (referred to as iteration 2 contribution) is added to the reconstructed profile from the first iteration to yield the overall profile of iteration 2 shown in figure 8(c). The process is then repeated. Assuming that reconstructed contributions from different iterations are independent random variables, covariance matrices from all iterations can be added up, yielding the cumulative estimated standard deviation plotted in figure 9(a).

Table 4. Summary of the performance parameters for the setups in figure 7.

| Setup | Total rms error | Local rms error | Processing time (seconds) |
|-------|----------------|----------------|--------------------------|
| $N_s = 30$ w/o TR | 0.01146 | 0.01122 | 36 |
| $N_s = 30$ w/ TR | 0.0306 | 0.0135 | 2 |
| $N_s = 10$ w/o TR | 0.0267 | 0.0315 | 5.4 |

We note several interesting points here. In the early iterations, the cumulative estimated error is not an accurate measure for the actual error; this is because of the deficiency of the underlying Born approximation in precisely modeling the scattered field at these stages. With increasing iterations, the discrepancy between actual and estimated errors gets smaller. As the method proceeds, the reconstructed contribution gets smaller and smaller, and so does the associated estimated standard deviation. However, the standard deviation decreases at a slower rate because the SNR of each inversion also decreases; this explains the increase in the percentile error shown in figure 9(b) with iterations. The percentile error is inversely proportional to the confidence level; therefore, a maximum threshold can be set on the former to determine when to stop. Intuitively, the higher SNR we have, the further we can go on with iterations, and the more accurate the inversion will be for a given (desired) confidence level.
For comparison, the convergence curve of a conventional (non-Bayesian) DBIM with a conjugate gradient (CG) solver is shown in figure 9(a). Besides being able to provide the confidence level of the inversion after each iteration, BDBIM shows faster convergence. This is because in conventional DBIM, the error signal (difference between true medium and synthetic responses) gets progressively overwhelmed by the measurement noise as the iterations proceed, whereas BCS inversion avoids over-fitting noise through the employed

Figure 8. Bayesian DBIM. (a) Forward permittivity distribution. (b)-(f) Reconstructed profiles from five iterations. \(N_s = 30\) and SNR = 10 dB.

Figure 9. Error analysis of Bayesian DBIM. (a) Actual total error and cumulative standard deviation after each iteration. For comparison, actual error of conventional DBIM with CG solver is also shown. (b) Estimated standard deviation of individual contributions from each iteration. Percentile error is the ratio of the estimated standard deviation to the rms of the differential contrast function contributed from each iteration.
sparsity priors. This observation is consistent with the results presented in [28, 29] for sparse scatterers.

7. Layered and structured media examples

Two examples of layered media are shown in figure 10. In figure 10(a), CH sensors are used to reconstruct a slanted layered medium with abrupt changes in permittivity. The problem is solved as a 2D problem and the reconstructed profile is shown in figure 10(b). Figure 10(c) shows a quasi-horizontally layered medium. It is a realization of a continuous random Gaussian medium with $l_{cx} = 125$ m along the $x$-direction and $l_{cy} = 1.25$ m along the $y$-direction. This is a good model for layered Earth formations encountered in geophysical exploration [60–62]. Prior knowledge of the layered nature of a problem can simplify the inversion significantly by solving the problem as a 1D inversion problem (i.e., restricting the unknowns to spatial harmonics along the $y$-direction), as shown in figure 10(d) for a BH scenario. The linear array shown in figure 10(c) can be deployed horizontally along the
$x$-direction as a surface controlled source electromagnetic (CSEM) array. In that case, the problem becomes 1D along a direction normal to the array. Finally, an example of a structured medium composed of piecewise-constant contrast objects is shown in figures 11. The inversion results shown in figures 10(b) and 11(b) illustrate that the proposed methodology can also be used for reconstructing structured media (i.e., not necessarily randomly distributed according to a Gaussian distribution). Note that the results shown correspond to a single-iteration inversion and that, if needed, BDBIM can be used for higher contrasts.

8. Conclusion

An approach based on Bayesian compressive sensing (in a broad sense of the term) was applied for ultrawideband multistatic inverse scattering problems in continuous random as well as structured media. It was shown that UWB BCS not only provides accurate reconstructions in different scenarios but also provides, a means for estimating the accuracy of the inversion. In addition, it allows for a systematic way of determining optimal locations so that the information gain is maximized by using sequential measurements. Furthermore, time-reversal-based focusing was combined with UWB BCS to achieve a localized, adaptive inversion and reduce overall inversion costs. The proposed methodology was successfully applied to a number of canonical geophysical imaging problems.

In principle, the proposed BCS inversion can be applied to all scenarios where DBIM is applicable, even in scenarios where the unknown is not truly sparse, like the structured media examples shown in section 7, if a sufficient number of measurements is conducted. Bayesian compressive sensing inversion was shown to have faster convergence than conventional DBIM. Similarly to other inversion techniques, the resolution of BCS inversion depends on the angular span of the acquisition (e.g., a full-aspect array provides better resolution than a linear array), the bandwidth (the highest resolvable spatial harmonic is in the same order of magnitude as the highest wavenumber), and the losses in the medium (higher losses result in less penetration, which limits the maximum frequency that can be used). Finally, we note that frequency dispersion in the medium has been neglected in the examples considered in this
paper. Application of BCS inversion to frequency-dispersive media will be considered in a future work.

Acknowledgements

This work has been supported by the National Science Foundation (NSF) under grant ECCS-0925272, by the Ohio Supercomputing Center (OSC) under grants PAS-0061 and PAS-0110, and by an Ohio State University Presidential Fellowship. We thank Dr L Carin and his group at Duke University for making the RVM code available online.

References

[1] Gerstoft P and Mecklenbrauker C F 1998 Ocean acoustic inversion with estimation of a posteriori probability distributions J. Acoust. Soc. Am. 104 808–19
[2] Birsan M 2003 A Bayesian approach to electromagnetic sounding in a marine environment IEEE Trans. Geosci. Remote Sensing 41 1455–60
[3] Ernst J R, Maurer H, Green A G and Holliger K 2007 Full-waveform inversion of crosshole radar data based on 2D finite-difference time domain solutions of Maxwells equations IEEE Trans. Geosci. Remote Sensing 45 2807–28
[4] Meles G A, der Kruk J V, Greenhalgh S A, Ernst J R, Maurer H and Green A G 2010 A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data IEEE Trans. Geosci. Remote Sensing 48 3391–407
[5] Zhang T, Chaumet P C, Mudry E, Sentenac A and Belkebir K 2012 Electromagnetic wave imaging of targets buried in a cluttered medium using a hybrid inversion–DORT method Inv. Problems 28 125008
[6] Dennison M and Devaney A J 2003 Inverse scattering in inhomogeneous background media Inv. Prob. 19 855–70
[7] Ybarra G A and Liu Q H 2008 Emerging Technologies in Breast Imaging and Mammography (Valencia, CA: American Scientific) chapter 16 Breast imaging using electrical impedance tomography
[8] Hassan A and El-Shenawee M 2011 Review of electromagnetic techniques for breast cancer detection IEEE Rev. Biomed. Eng. 4 103–18
[9] Schmidt D, George J and Wood C 1999 Bayesian inference applied to the electromagnetic inverse problem Hum. Brain Map. 7 195–212
[10] Gilmore C, Abubakar A, Hu W, Habashy T M and van der Berg P M 2009 Microwave biomedical data inversion using the finite-difference contrast source inversion method IEEE Trans. Antennas Propag. 57 1528–38
[11] Winters D W, Shea J D, Madsen E L, Frank G R, Veen B D V and Hagness S C 2008 Estimating the breast surface using UWB microwave monostatic backscatter measurements IEEE Trans. Biomed. Eng. 55 247–56
[12] Davis S K, Veen B D V, Hagness S C and Kelcz F 2008 Breast tumor characterization based on ultrawideband microwave backscatter IEEE Trans. Biomed. Eng. 55 237–46
[13] Baranoski E J 2008 Through-wall imaging: historical perspective and future directions J. Franklin Inst. 345 556–69
[14] Attiya A M, Bayram A, Safaai-Jazi A and Riad S M 2004 UWB applications for through-wall detection Proc. IEEE Antennas Propag. Int. Symp. 3 3079–82
[15] Li L, Zhang W and Li F 2010 A novel autofocus approach for real-time through-wall imaging under unknown wall characteristics IEEE Trans. Geosci. Remote Sensing 48 423–431
[16] Takagi T, Bowler J R and Yoshida Y 1997 Electromagnetic Nondestructive Evaluation (Amsterdam: IOS)
[17] Lasri T and Zoughi R (ed) 2001 Advances and applications in microwave and millimeter wave nondestructive evaluation Subsurface Sensing Technol. Appl. 2 (special issue)
[18] Lesselier D and Bowler J (ed) 2002 Electromagnetic and ultrasonic nondestructive evaluation Inv. Problems 18 (special issue)
[19] Leuschen C J and Plumb R G 2006 A matched-filter-based reverse-time migration algorithm for ground-penetrating radar data IEEE Trans. Geosci. Remote Sensing 39 1257–64
[20] Foroozan F and Asif A 2010 Time-reversal ground-penetrating radar: range estimation with crerram lower bounds IEEE Trans. Geosci. Remote Sensing 48 3698–3708
[21] Xu X, Miller E L, Rappaport C M and Sower G D 2002 Statistical method to detect subsurface objects using array ground-penetrating radar data IEEE Trans. Geosci. Remote Sensing 40 963–76
[22] Potin D, Duflos E and Vanheeghe P 2006 Landmines ground-penetrating radar signal enhancement by digital filtering IEEE Trans. Geosci. Remote Sensing 44 2393–406
[23] Li M, Abubakar A and Habashy T M 2010 Application of a two-and-a-half dimensional model-based algorithm to crosswell electromagnetic data inversion Inv. Problems 26 074013
[24] Spies B and Oristaglio M (ed) 1999 Three-dimensional Electromagnetics (Tulsa, OK: Society Of Exploration Geophysicists)
[25] Tarantola A 2005 Inverse Problem Theory and Methods for Model Parameter Estimation (Philadelphia, PA: SIAM)
[26] Idier J 2010 Bayesian Approach to Inverse Problems (Hoboken, NJ: Wiley)
[27] Lemm J C 2003 Bayesian Field Theory (Baltimore, MD: Johns Hopkins University Press)
[28] Oliveri G, Rocca P and Massa A 2011 A Bayesian-compressive-sampling-based inversion for imaging sparse scatterers IEEE Trans. Geosci. Remote Sensing 49 3993–4006
[29] Poli L, Oliveri G and Massa A 2012 Microwave imaging within the first-order Born approximation by means of the contrast-field Bayesian compressive sensing IEEE Trans. Antennas Propag. 60 2865–79
[30] Chaloner K and Verdinelli I 1995 Bayesian experimental design: a review Stat. Sci. 10 237–304
[31] Ji S, Xue Y and Carin L 2008 Bayesian compressive sensing IEEE Trans. Signal Process. 56 2346–56
[32] Ji S, Dunson D and Carin L 2009 Multi-task compressive sensing IEEE Trans. Signal Process. 57 92–106
[33] Tipping M E 2001 Sparse Bayesian learning and the relevance vector machine J. Machine Learning Res. 1 211–44
[34] Tipping M E and Faul A C 2003 Fast marginal likelihood maximisation for sparse Bayesian models Proc. 9th Int. Workshop Artificial Intelligence and Statistics (KeyWest, FL, Jan. 3–6 2003)
[35] Marengo E A, Hernandez R D, Citron Y R, Gruber F K, Zambrano M and Lev-Ari H 2008 Compressive sensing for inverse scattering Proc. 29th URSI Gen. Assem. (Jian. 7–16 2008)
[36] Marengo E A 2008 Compressive sensing and signal subspace methods for inverse scattering including multiple scattering Proc. IEEE Geoscience Remote Sensing Symp. (Jul. 7–11 2008)
[37] Marengo E A 2008 Subspace and Bayesian compressive sensing methods in imaging Proc. Prog. Electronagn. Res. Symp. (Jul. 2–6 2008)
[38] Mostofi Y 2013 Cooperative wireless-based obstacle/object mapping and see-through capabilities in robotic networks IEEE Trans. Mobile Comput. 12 817–29
[39] González-Ruiz A and Mostofi Y 2013 Cooperative robotic structure mapping using wireless measurements—a comparison of random and coordinated sampling patterns IEEE Sensors J. 13 2571–80
[40] Gonzalez-Ruiz A, Ghaffarkhah A and Mostofi Y 2014 An integrated framework for obstacle mapping with see-through capabilities using laser and wireless channel measurements IEEE Sensors J. 14 25–38
[41] Harada H, Wall D J, Takenaka T and Tanaka T 1995 Conjugate gradient method applied to inverse scattering problems IEEE Trans. Antennas Propag. 43 784–92
[42] Moghadam M and Chew W 1992 Nonlinear two-dimensional velocity profile inversion using time domain data IEEE Trans. Geosci. Remote Sensing 30 147–56
[43] Weedon W and Chew W 1993 Time-domain inverse scattering using the local shape function (LSF) method Inv. Problems 9 551–64
[44] Mora P 1987 Nonlinear two-dimensional elastic inversion of multi-offset seismic data Geophys. 52 1211–28
[45] Wang Y and Chew W 1990 Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method IEEE Trans. Med. Imag. 9 218–25
[46] Gelfand A E and Smith A F M 1990 Sampling-based approaches to calculating marginal densities J. Am. Stat. Assoc. 85 398–409
[47] Walsh B 2004 Markov chain Monte Carlo and Gibbs sampling http://web.mit.edu/~wingated/www/introductions/mcmc-gibbs-intro.pdf.
[48] Fouda A E and Teixeira F L 2014 Statistical stability of ultrawideband time-reversal imaging in random media IEEE Trans. Geosci. Remote Sensing 52 87–9
[49] Fouda A E, Lopez-Castellanos V and Teixeira F L 2014 Experimental demonstration of statistical stability in ultrawideband time-reversal imaging IEEE Geosci. Remote Sensing Lett. 11 29–33
[50] Candes E J and Wakin M B 2008 An introduction to compressive sampling IEEE Signal Process. Mag. 25 21–30
[51] Potter L C, Ertin E, Parker J T and Cetin M 2010 Sparsity and compressed sensing in radar imaging Proc. IEEE 98 1006–20
[52] Taflove A and Hagness S 2005 Computational Electrodynamics: The Finite-difference Time-domain Method 3rd edn (Norwood, MA: Artech House)
[53] Cover T M and Thomas J A 1991 Elements of Information Theory (New York: Wiley)
[54] Fink M 1993 Time reversal mirrors J. Phys. D.: Appl. Phys. 26 1333–50
[55] Yavuz M E and Teixeira F L 2008 On the sensitivity of time-reversal imaging techniques to model perturbations IEEE Trans. Antennas Propagat. 56 834–43
[56] Fouda A E and Teixeira F L 2012 Imaging and tracking of targets in clutter using differential time-reversal techniques Waves Random Complex Media 22 66–108
[57] Liu D, Vasudevan S, Krolik J, Bal G and Carin L 2007 Electromagnetic time-reversal source localization in changing media: experiment and analysis IEEE Trans. Antennas Propag. 55 344–54
[58] Moghadam M, Chew W and Oristaglio M 1991 Comparison of the Born iterative method and Tarantola’s method for an electromagnetic time-domain inverse problem Int. J. Imag. Syst. Technol. 3 318–33
[59] Wang Y and Chew W 1989 An iterative solution of the two-dimensional electromagnetic inverse scattering problem Int. J. Imag. Syst. Technol. 1 100–8
[60] Liu G-S, Teixeira F L and Zhang G-J 2012 Analysis of directional logging tools in anisotropic and multieccentric cylindrically layered Earth formations IEEE Trans. Antennas Propag. 60 318–27
[61] Hue Y-K and Teixeira F L 2007 Numerical mode-matching method for tilted coil antennas in cylindrically layered anisotropic media with multiple horizontal beds IEEE Trans. Geosci. Remote Sensing 45 2451–62
[62] Lee H O and Teixeira F L 2007 Cylindrical FDTD analysis of LWD tools through anisotropic dipping-layered Earth media IEEE Trans. Geosci. Remote Sensing 45 383–8