

Subsurface Sensing Under Sensor Positional Uncertainty

Ashley B. Tarokh, *Member, IEEE*, and Eric L. Miller, *Senior Member, IEEE*

Abstract—We consider the problem of classifying buried objects using electromagnetic induction data collected in a setting where there are errors in sensor positioning. Using a series of decay constants (or equivalently, Laplace plane poles) as features for classification, our algorithm seeks to estimate these poles and, subsequently, to determine the type of object in the sensor field of view. In many practical scenarios, a set of data is often accompanied by domain knowledge that the location of the transmitters and/or receivers is only known to within some degree of accuracy (e.g., 10 cm in the along-track direction and 5-cm cross-track). Here, we develop an approach to the extraction of information from such data sets in which the quantitative positional bound information is used in the context of a min-max optimization strategy. Specifically, we look for the parameters of interest that minimize the maximum data residual, where the maximum error is computed over ellipsoids or polyhedra of possible sensor locations defined by the bound information. Our formulation admits data collection with independent or dependent positional uncertainty values at successive nominal collection locations. Our algorithms for solving this optimization problem are validated using simulated and measured data.

Index Terms—Bounded-data uncertainty, dynamic programming, electromagnetic induction (EMI), min-max optimization, object classification, parameter estimation, positional uncertainty, subsurface sensing.

I. INTRODUCTION

IN RECENT years, there has been significant effort toward the development of physics-based signal and image-processing methods for the solution of subsurface-sensing problems given the data from one or more sensors that have been scanned over a region of the Earth [1]–[16]. Broadly speaking, such algorithms either develop a voxelated image of the subsurface or directly extract characteristics (size, shape, location, class, etc.) from the data of buried objects such as unexploded ordnance, landmines, and utilities [17], [18]. In either case, the parameters of interest are typically chosen using an algorithm that minimizes some measure of the error between the observed data and the predictions of a computational sensor model.

Manuscript received April 20, 2006; revised September 19, 2006. This work was supported in part by the Strategic Environmental Research and Development Program under projects CU-1217 and MM-1379 and in part by the National Science Foundation under Grant 0208548.

A. B. Tarokh is with the Center for Bioinformatics, Harvard Center for Neurodegeneration and Repair, Functional and Molecular Imaging Center, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115 USA.

E. L. Miller is with the Department of Electrical and Computer Engineering, Tufts University, Medford, MA 02155 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2006.888851

A tacit assumption in almost all of these methods is that the positions of the sensor are known precisely relative to some fixed reference point, a condition that is often not met in practice [14], [15], [19]. For handheld systems especially, the sensor may not be equipped with global positioning system (GPS), in which case some sort of dead reckoning may be employed to get approximate locations [20]. In the case of vehicular or cart-mounted sensors, even with GPS, the effects of positional uncertainties have not been extensively studied for problems requiring high-resolution localization of buried objects. In [14], Tatum *et al.* propose several approaches to mitigate sensor positional uncertainty including low-pass filter and hidden Markov-model techniques to correct for positional errors. In [13], signal-processing techniques for data obtained from multi-axis sensors were proposed and studied using simulated sensor-positional uncertainties. A report on a workshop regarding unexploded-ordnance (UXO) detection methods is presented in [19]. This workshop raised many questions regarding state-of-the-art UXO detection methods, among which was the issue of sensor positional uncertainty.

The goal of the work in this paper is the development and analysis of a general-purpose approach to these inverse problems, which explicitly accounts for positional uncertainty. Our model considers the classification problem, where the features of interest are given by a set of decay constants (or equivalently, Laplace plane poles). However, our technique applies more generally and need not be restricted to this particular classification problem. The method we propose is motivated by the observation that real data sets for which positions of the sensor are recorded are often accompanied by the caveat that these locations are accurate only to within some tolerances in the x , y , and z coordinates.

Parameter estimation in the presence of uncertainty has historically been the subject of intensive study [21]–[32]. This uncertainty can either be due to the presence of noise in the data or due to the existence of uncertainty in the model [30]. In each case, we can handle the uncertainty either deterministically or stochastically. Least squares (LS) [29] and ridge regression [24], [33], [34] are two well-known deterministic models used in the literature for handling uncertainties caused by the presence of noise in the data. The linear LS criterion was developed by Gauss in the 1700s. This method is very attractive, because it results in a closed-form solution that can be updated, as more data become available.

When dealing with uncertainties in the model, deterministic measures such as robust estimation and filtering techniques (or H^∞) [22], [25], [26], [28], [31], [32] and total LS (TLS) [23], [27] have been used. The H^∞ method is essentially an

optimization technique designed to minimize the worst case estimation error [22], [26]. The TLS method is an alternative method to LS for obtaining a solution to linear problems [23]. This method allows one to incorporate errors into the data matrix when formulating the problem. However, in cases where the effects of noise can be overemphasized, this method tends to result in overconservative estimates [21]. The authors of [21] propose an approach that avoids this overcompensation by employing a method that is very similar to H^∞ . However, [21] uses a cost function that is more complex than H^∞ and, in contrast to robust estimation, imposes constraints on the error values. A problem similar to that of TLS arises when using H^∞ filtering. Since the design of such problems is such that they are robust to uncertainties in the data, the solutions may again lead to overcompensated results [21], [23].

Among stochastic methods designed to deal with the existence of uncertainties in the data, the predominant methods are maximum *a posteriori* and maximum likelihood [35]. These two methods often result in similar mathematical solutions with varying interpretations for the components of the algorithm. One stochastic approach proposed for dealing with uncertainties in the model is by directly incorporating this information into a processing scheme that is based on Bayesian statistics [5], [36]–[38]. Specifically, the perturbations in the positions of the sensor are modeled as random variables. Typically, these random variables are taken to be independent and, in the absence of any other information, uniformly distributed between bounds related to the tolerances inherent in the measurement device. The resulting statistical estimation or decision problem is then solved via Monte Carlo integration methods to “average away” the effects of these unknowns [36].

Here, we explore an approach to dealing with uncertainty in sensor location that does not require the explicit specification of a probability density for these quantities but rather is motivated by practical application, in which practitioners quantify the accuracy of their sensor positioning through the use of bounds, e.g., along-track positioning is accurate to within ± 5 cm. In these circumstances, we think it is useful to exploit this bound information in the parameter-estimation portion of the processing using not a stochastic-processing approach but rather a “min–max” formulation of the estimation problem. That is, we look for those parameters (pole values, rotation angles, etc.) which minimize the maximum of some cost function, where the maximum has taken over all possible sensor locations, each of which is restricted to lie within some bounded region of space, where the size of the region is dictated by the accuracy bounds discussed above. To the best of our knowledge, a min–max formulation of the problem of uncertain sensor locations has not been proposed or examined to date in the context of subsurface sensing. Thus, our effort provides a new and potentially useful way of approaching this issue. We feel, moreover, that is, in a sense, a very natural way to deal with the bound information often associated with sensor positioning for UXO-type data sets.

Our approach is divided into two possible scenarios of interest. First, we consider the case where the positional uncertainties at successive data-collection locations are independent. This models a situation where data-collection locations are

selected with reference to some absolute position. We analyze polyhedral and ellipsoidally shaped uncertainty regions and arrive at very tractable and potentially easily parallelized algorithms for solving the resulting min–max optimization problem. This model, however, does not admit dependent positional uncertainties, which can occur as a result of an accumulation of positional uncertainty as the sensor moves over the data-collection locations.

Under dependent positional uncertainties, we detail a dynamic-programming (DP) formulation as a means of searching among the possible sequences of data-collection locations. This formulation also admits polyhedral and ellipsoidally shaped uncertainty regions. Although it is more computationally intensive than the case of independent errors, the results we see from the simulation are promising and point to the utility of further work involving much more tractable approximate dynamic-programming approaches.

We report the results of numerical experiments involving both simulated and field data and compare these results to those of an existing technique that does not account for positional uncertainty. In all cases, we observe significant classification-performance improvements compared to cases where sensor-positioning errors are not taken into account.

The remainder of this paper is organized as follows. In Section II, we introduce the classification problem of interest and develop the positional-uncertainty framework. We consider the case of independent positional uncertainties in Section III. For dependent positional uncertainties, a dynamic-programming formulation is presented in Section III-D. Section IV reports the results of our numerical studies for UXO characterization using electromagnetic induction (EMI) sensors. Finally, conclusions and future work are provided in Section V.

II. GENERAL PROBLEM FORMULATION

A. Classification Based on Spatially Sampled EMI Data

We consider the problem of classifying objects based on spatially sampled EMI data in the time or frequency domain. Data are assumed to be collected from a single monostatic sensor, as depicted in Fig. 1. The generalizations to multiple sets of data from multistatic-sensing systems are conceptually straightforward but notationally burdensome. It is assumed that our sensor stops to collect data at N distinct-transmitter/receiver-location combinations. We associate with the i th data-collection location the nominal coordinate $\mathbf{r}_i = (x_i, y_i, z_i)$, for $i = 1, \dots, N$. At each location, we assume that M time or frequency samples are taken. Under this model, the k th data sample at the j th sensor location is given by [17]

$$d_{j,k} = g_j^T R^T \Lambda_k R f_j + \sigma n_{j,k} \quad (1)$$

$$= s_{j,k} + \sigma n_{j,k} \quad (2)$$

where g_j is a 3×1 vector holding the (x , y , and z) components of the magnetic field induced at location \mathbf{r}_j by a current of I units flowing through the receive coil, and f_j represents the excitation-field vector evaluated at the dipole position. Functional forms for f and g are provided in [39]. The noise variance

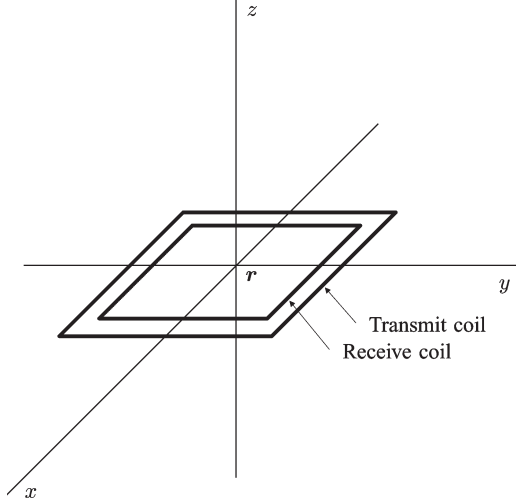


Fig. 1. Monostatic sensor for data collection, consisting of two concentric coils for transmitting and receiving electromagnetic energy. The sensor is centered at nominal location \mathbf{r} .

is σ and $n_{j,k}$ is a zero-mean unit-variance white Gaussian random variable used to represent the measurement noise. The 3×3 matrix R is a 3-D rotation matrix used to transform the global-coordinate frame of reference to the rotated local frame of reference at the object. Algebraically, the matrix R diagonalizes the complex-valued magnetic polarizability tensor Λ_k . R is parameterized by three Euler angles ψ_1, ψ_2 , and ψ_3 and is modeled under the x -convention [40]. Thus, it is without loss of generality that we assume Λ_k is a diagonal 3×3 complex-valued matrix. For the frequency-domain version of the problem, Λ_k has the form

$$\Lambda(\omega) = \begin{bmatrix} \lambda_1(\omega) & & \\ & \lambda_2(\omega) & \\ & & \lambda_3(\omega) \end{bmatrix}. \quad (3)$$

Replacing ω by t in (3) yields the time-domain form of Λ_k . The quantities λ_1, λ_2 , and λ_3 are associated with one of each of the principle axes of the object. In other words, the diagonal elements of Λ_k hold the scattering characteristics of the object along each of the three principle axes for the k th frequency (time) sample. The model for these quantities is given as follows [41]:

$$\lambda_i(\omega) = \sum_{l=1}^{\infty} \frac{a_{i,l} j\omega}{p_{i,l} + j\omega}, \quad i = 1, 2, 3 \quad (4)$$

$$\lambda_i(t) = - \sum_{l=1}^{\infty} a_{i,l} p_{i,l} e^{-p_{i,l} t} u(t), \quad i = 1, 2, 3 \quad (5)$$

where $j = \sqrt{-1}$, $a_{i,l}$ is the expansion coefficient for the l th term corresponding to the i th axis, $p_{i,l}$ is the l th pole for the i th axis, and $u(t)$ is the unit-step function. Equations (4) and (5) only hold for nonferrous objects. In the case of ferrous objects, a dc offset must be added in the frequency-domain version of the model. Correspondingly, a Dirac delta function must be included in the time-domain version of the model.

To show the explicit dependence of our model on both the object parameters and the positions of the sensors at each data-collection location, we express, using (2), the model for the k th data value collected at the j th location as

$$s_k(\boldsymbol{\theta}, \mathbf{r}_j) = g^T(\boldsymbol{\theta}, \mathbf{r}_j) R^T(\boldsymbol{\theta}) \Lambda_k(\boldsymbol{\theta}) R^T(\boldsymbol{\theta}) f(\boldsymbol{\theta}, \mathbf{r}_j) \quad (6)$$

where $\boldsymbol{\theta}$ is a 6-D vector that holds the coordinates of the object (x, y , and z) as well as the Euler angles (ψ_1, ψ_2 , and ψ_3).

The classification approach is started by constructing a target-signature library, which will be used in the actual processing. For each target of interest, this library will hold the corresponding pole values [17]. Given the library, classification is a two-step process. First, for each target in the library, the data $d_{j,k}, j = 1, \dots, N$ and $k = 1, \dots, M$ are used to estimate the unknown parameters associated with that model: poles, expansion coefficients, object location, and object orientation. Second, using these estimates, we perform the classification. The classifier used in this paper is based on the fit of the k th model in the library to the available data. This involves evaluating (6) at the estimated parameter values (poles, expansion coefficients, object location, and object orientation) for the object, obtained in the first step. As explained and demonstrated more fully in [17], this approach is likely to be of use when the signal-to-noise ratio (SNR) is relatively low. It is important to note that we only study the databased-classification method as a proof-of-concept. Other classification techniques, for example, those explored in [17], can also be applied without requiring any changes to our proposed methods. Additionally, it is often of interest to incorporate clutter rejection into the classifier. Such an extension can be easily incorporated into a classifier by adopting a thresholding technique (see, for example, [18] and references therein).

B. Including Positional Uncertainty in the Model

As we motivated in Section I, it is often the case that the precise sensor position at each sensor location is not known. To model this situation, the precise coordinates of the j th location \mathbf{r}_j are taken to be the sum of two components, for each $j = 1, \dots, N$

$$\mathbf{r}_j = \mathbf{r}_{0,j} + \delta \mathbf{r}_j \quad (7)$$

where $\mathbf{r}_{0,j}$ is the nominal (equivalently expected or understood) position for the j th sensor location and $\delta \mathbf{r}_j$ is the perturbation to $\mathbf{r}_{0,j}$ that yields the true sensor position at the j th data-collection location. We denote the region of uncertainty within which the true i th sensor location is known to reside as \mathcal{S}'_j . The set of error tolerances to which the nominal-location information is subjected is defined for the j th location by

$$\mathcal{S}_j = \{ \delta \mathbf{r} = \mathbf{r} - \mathbf{r}_{0,j} : \mathbf{r} \in \mathcal{S}'_j \}. \quad (8)$$

We assume that \mathcal{S}_j is a connected convex and bounded region for each $j = 1, \dots, N$, for example, a box or an ellipsoid. This region could be a conservative estimate of the true set of error

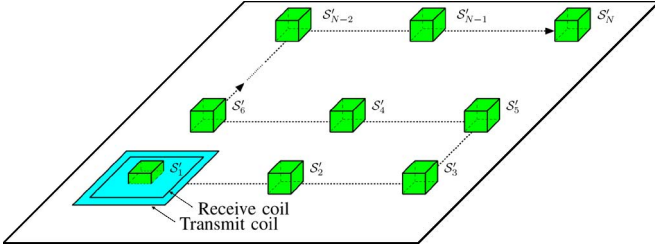


Fig. 2. Data are collected at N locations over a field of interest. Location j has associated with it a region of uncertainty \mathcal{S}'_j , shown here for every j as a box-shaped region. The dashed line connecting data-collection locations indicates the sequence in which data are collected over the field.

tolerances in measurement but reflects the best range that is known. Fig. 2 depicts the sensor locations over which data are collected. Note that the region of uncertainty is depicted at each location as a box.

Fig. 2 depicts the successive sensor positions as being visited sequentially. Sequential data collection is common; modeling data collection that relies upon dead reckoning (for example, a device employing an inertial measurement unit). In such a scenario, it is the case that successive perturbation values are dependent, since each successive location is determined nominally from the last location, but there is some uncertainty in the calculation of the new position. Given the manner in which data are collected [20], it is reasonable to assume that successive data-collection locations depend only the current sensor positions and not on past sensor positions. In probability terms, the set of sensor positions evolves in a Markovian fashion [42]. Consequently, this scenario is effectively modeled through a first-order autoregressive relationship with parameter α as follows. Suppose that nominal locations $\mathbf{r}_1, \dots, \mathbf{r}_N$ are fixed values corresponding to the known but inexact sensor locations. Then, the sets $\mathcal{S}_1, \dots, \mathcal{S}_N$ are related according to

$$\mathcal{S}_j(\delta\mathbf{r}_{j-1}) = \{\delta\mathbf{r}_j = \alpha\delta\mathbf{r}_{j-1} + \delta\mathbf{r} : \delta\mathbf{r} \in \mathcal{S}_j^{\text{new}}\} \quad (9)$$

for $j = 2, \dots, N$. We have explicitly shown the dependence of the j th uncertainty region on the perturbation at the $(j-1)$ th sensor location by defining \mathcal{S}_j as a set-valued function. In (9), we refer to the set $\mathcal{S}_j^{\text{new}}$ as the set of innovations at the j th location. This set reflects the new positional uncertainty introduced by measurement at the j th location. In words, our model for successive uncertainty regions begins with the first uncertainty region \mathcal{S}_1 . The next uncertainty region \mathcal{S}_2 is defined in terms of an element of \mathcal{S}_1 . Hence, given that some uncertainty $\delta\mathbf{r}_1$ was realized in the region \mathcal{S}_1 , the uncertainty value at the second sensor location $\delta\mathbf{r}_2$ must depend on $\delta\mathbf{r}_1$. The actual dependence is given by $\delta\mathbf{r}_2 = \alpha\delta\mathbf{r}_1 + \delta\mathbf{r}$, where $\delta\mathbf{r}$ is the additional uncertainty (or innovation) that compounds with the initial uncertainty. This new uncertainty is restricted to the set $\mathcal{S}_j^{\text{new}}$. The same relationship follows for subsequent sensor locations.

The parameter α essentially captures the degree to which the sensor positional uncertainty of the previous data-collection location affects that of the subsequent data-collection location.

Note that when $\alpha = 0$, this model specializes to the case where each perturbation term is independent of the others, which is a common scenario when a fixed reference (such as GPS) is used for calculation of each sensor location [14]. A nonzero value of α can arise under several circumstances, such as when no fixed reference is available and, consequently, error is compounded due to fluctuations or poor estimates of the velocity of the sensor as it collects data at regularly spaced intervals of time or due to miscalculations from equipment or human failure. The value $\alpha = 1$ is particularly applicable to data-collection scenarios where dead reckoning with no fixed reference is available. In such a situation, the current sensor position is treated as the current nominal data-collection location when computing the subsequent data-collection location. The value of α clearly depends on the data-collection mechanism and must be modeled appropriately in order to reach the best possible optimization goals.

We denote by \mathcal{R} the set of admissible trajectories of perturbation values, $(\delta\mathbf{r}_1, \dots, \delta\mathbf{r}_N)$ under the above model (9). \mathcal{R} may be expressed as

$$\mathcal{R} = \{(\delta\mathbf{r}_1, \dots, \delta\mathbf{r}_N) : \delta\mathbf{r}_1 \in \mathcal{S}_1, \delta\mathbf{r}_j \in \mathcal{S}_j(\delta\mathbf{r}_{j-1}) \quad \forall j = 2, \dots, N\}. \quad (10)$$

The set of admissible true sensor locations is then given by

$$\mathcal{R}' = \{(\mathbf{r}_1, \dots, \mathbf{r}_N) : \mathbf{r}_j = \mathbf{r}_{0,j} + \delta\mathbf{r}_j \quad \forall j, (\delta\mathbf{r}_1, \dots, \delta\mathbf{r}_N) \in \mathcal{R}\}. \quad (11)$$

III. MIN-MAX FORMULATION FOR UNKNOWN BUT BOUNDED POSITIONAL UNCERTAINTY

In this section, we formulate a min-max optimization strategy for our classification problem, which is subjected to sensor positional uncertainty. Our mathematical formulation requires, first, a cost function over which to optimize. As is typical in many data inversion applications where measurement noise is modeled using a white Gaussian noise process, here, we employ a cost functional equal to the norm of the squared-error difference between the data observed and the prediction of the data provided by the model \mathbf{s} over all $\boldsymbol{\theta}$

$$\hat{J}_1(\boldsymbol{\theta}, \mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{j=1}^N \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_j)\|^2. \quad (12)$$

The problem now is how to specify the \mathbf{r}_j in function s_k in \hat{J}_1 . Typically, the \mathbf{r}_j are set equal to their nominal values $\mathbf{r}_{0,j}$, and an LS estimate of $\boldsymbol{\theta}$ is found. Here, we seek to augment the minimum error notion of optimality in a manner that easily incorporates the additional knowledge we have concerning the geometry of the \mathcal{S}'_j for $j = 1, \dots, N$. Essentially, the problem we have is to determine a set of primary parameters $\boldsymbol{\theta}$ in the presence of a collection of “nuisance” parameters, the \mathbf{r}_j , which are restricted to exist in known regions of space. This additional knowledge concerning the positional uncertainties may be

incorporated into our cost functional to yield the augmented cost functional \hat{J}_2 , given by¹

$$\hat{J}_2(\boldsymbol{\theta}, \mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{j=1}^N \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_j)\|^2 + \gamma \sum_{j=2}^N \|\delta \mathbf{r}_j - \alpha \delta \mathbf{r}_{j-1}\|^2. \quad (13)$$

The cost functional in (13) penalizes simultaneously the fidelity of the data reconstruction implied by the parameter vector $\boldsymbol{\theta}$ and the nearness of adjacent perturbation values, related through the coefficient α . This last term is motivated by the recursively defined model we have hypothesized for the sets \mathcal{S}_j which implies a degree of nearness for the positional perturbations for successive data-acquisition locations.

The parameter γ in (13) may be considered as a regularization coefficient that establishes the level of tradeoff between the two terms in (13). The choice of appropriate γ is a well-studied issue in the field of inverse problems. Common approaches include the L-curve and generalized cross-validation methods [43], [44]. In our case where, as we discuss below, (13) is used in the context of a dynamic-programming algorithm, it is not at all clear that such regularization-parameter-selection methods would be applicable or appropriate. Developing such methods, while very interesting, is well beyond the scope of the work in this paper, where our primary concern is in demonstrating a new approach to the larger problem of processing in the presence of positional uncertainty. Hence, in the examples presented in Section IV, we chose γ through trial and error. In fact, the value of $\gamma = 1$ was used for all examples and provided a quite strong performance.

We note that (13) takes the form of the log-likelihood cost function that would result were we have to assume that the position errors themselves followed a first-order Gauss–Markov process with drift defined by α and driving noise variance related γ . We choose here not to use (13) as the basis for an extended-estimation problem where we jointly determine from the data values for both $\boldsymbol{\theta}$ as well as all of the $\delta \mathbf{r}_j$. For the UXO problem of interest here, $\boldsymbol{\theta}$ contains 13 parameters while the total number of positional parameters would be three times the number of sensor locations. Given specifically the well-known difficulties in estimating the decay rates for UXO objects [18], [36], [45], [46], the addition of this large number of nuisance parameters associated with the sensor locations would likely degrade the accuracy of these estimates quite severely.

In the following sections, we detail an alternative approach to processing, in which we explicitly make use of the bound information used to construct the sets \mathcal{S}_j . Specifically, we consider a “min–max” optimization problem based on (13), a well-known approach to problems involving auxiliary parameters whose values are not known exactly but are known to lie within a bounded region of space. This approach to the problem

amounts to selecting that $\boldsymbol{\theta}$ that minimizes the worst error, as measured by $\hat{J}_2(\boldsymbol{\theta}, \mathbf{r}_1, \dots, \mathbf{r}_N)$ as $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ ranges over \mathcal{R}' for $j = 1, \dots, N$. Formally, $\hat{\boldsymbol{\theta}}$, our estimate of $\boldsymbol{\theta}$, is defined through the min–max formulation as

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \max_{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathcal{R}'} \hat{J}_2(\boldsymbol{\theta}, \mathbf{r}_1, \dots, \mathbf{r}_N). \quad (14)$$

From (6), it is clear that s_k is nonlinear in the position values $\mathbf{r}_1, \dots, \mathbf{r}_N$. Since the fields for these sensors are very smooth relative to the spatial scales of the perturbation to the position, for sufficiently small values $\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N$, the following first-order approximation is valid:

$$s_k(\boldsymbol{\theta}, \mathbf{r}_j) = s_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) + \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) \delta \mathbf{r}_j + o(\delta \mathbf{r}_j) \quad (15)$$

where we have adopted the little-o notation to represent the higher order terms in the expansion of $s_k(\boldsymbol{\theta}, \mathbf{r}_j)$. Formally, if the function $h(\mathbf{r}) = o(\mathbf{r})$ then $(h(\mathbf{r})/\mathbf{r}) \rightarrow 0$ as $\mathbf{r} \rightarrow 0$ [47]. The quantity $\mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j})$ is a 1×3 vector, whose components, respectively, equal the partial derivative of s_k with respect to x_j, y_j , and z_j , evaluated at the respective nominal coordinates $x_{0,j}, y_{0,j}$, and $z_{0,j}$.

We may then apply (15) in (14) to yield a new min–max formulation with a cost functional that is quadratic in the unknown perturbations. Let

$$J(\boldsymbol{\theta}, \delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) = \sum_{j=1}^N \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) \delta \mathbf{r}_j\|^2 + \gamma \sum_{j=2}^N \|\delta \mathbf{r}_j - \alpha \delta \mathbf{r}_{j-1}\|^2 \quad (16)$$

$$= \|\bar{\mathbf{d}} - \bar{\mathbf{A}} \delta \mathbf{r}\|^2 \quad (17)$$

where the simple expression in (17) follows easily from (16) by appropriately defining vector $\bar{\mathbf{d}}$ and matrix $\bar{\mathbf{A}}$. Furthermore, we have represented $\delta \mathbf{r} = (\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N)$. Then, our simplified min–max formulation is given by

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \max_{(\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) \in \mathcal{R}} J(\boldsymbol{\theta}, \delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N). \quad (18)$$

Our next result allows us to conclude that the optimal value of the inner maximization in (18) is achieved on the boundary of \mathcal{R} .

Theorem 1: The vector $(\delta \mathbf{r}_1^*, \dots, \delta \mathbf{r}_N^*)$ achieving the maximum value in

$$\max_{(\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) \in \mathcal{R}} J(\boldsymbol{\theta}, \delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) \quad (19)$$

lies on the boundary of \mathcal{R} .

Proof: See Appendix A. ■

Applying Theorem 1, we can conclude that the space of feasible vectors $(\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N)$, over which the inner maximization in (18) must be taken, can be greatly reduced by searching over only the boundary of \mathcal{R} . We may not, however, conclude that $\delta \mathbf{r}_j$ lies on the boundary of \mathcal{S}_j for $j = 1, \dots, N$. We shall

¹The functional \hat{J}_2 does not show an explicit dependence on the perturbation values $\delta \mathbf{r}_j$, $j = 1, \dots, N$, since these are given through (7), using the known nominal values $\mathbf{r}_{0,j}$ and $j = 1, \dots, N$.

demonstrate (in Section III-A) that in the case of $\alpha = 0$, this property is indeed true and results in a great simplification of the problem. For general $\alpha > 0$, we will employ a dynamic-programming algorithm to obtain the optimal inner maximization. This general formulation appears in Section III-D.

A. Case of $\alpha = 0$: Separable Optimization

When $\alpha = 0$, there is no assumed or known relationship between the adjacent perturbation values. Consequently, the penalty incorporated in the cost functional J associated with the relationship between adjacent perturbations is treated in this section as being unnecessary. We ensure this is the case by maintaining $\gamma = 0$, whenever $\alpha = 0$. Thus, for $\alpha = 0$, we have the following separation property:

$$J(\boldsymbol{\theta}, \delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) = \sum_{j=1}^N \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) \delta \mathbf{r}_j\|^2 \quad (20)$$

$$= \sum_{j=1}^N \|\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j\|^2 \quad (21)$$

$$= \sum_{j=1}^N J_j(\boldsymbol{\theta}, \delta \mathbf{r}_j) \quad (22)$$

where we define $J_j = \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) \delta \mathbf{r}_j\|^2$, for $j = 1, \dots, N$. Under this separation, our optimization (18) can be expressed as

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \max_{(\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) \in \mathcal{R}} J(\boldsymbol{\theta}, \mathbf{r}_1, \dots, \mathbf{r}_N) \quad (23)$$

$$= \sum_{j=1}^N \arg \min_{\boldsymbol{\theta}} \max_{\delta \mathbf{r}_j \in \mathcal{S}_j} J_j(\boldsymbol{\theta}, \delta \mathbf{r}_j). \quad (24)$$

Since $\mathcal{S}_j(\delta \mathbf{r}_{j-1})$ has no dependence on $\delta \mathbf{r}_{j-1}$ when $\alpha = 0$, we have that $\mathcal{S}_j(\delta \mathbf{r}_{j-1}) = \mathcal{S}_j^{\text{new}}$ for $j = 2, \dots, N$. To simplify the notation, we have dropped the expression of \mathcal{S}_j as a function of $\delta \mathbf{r}_{j-1}$ for $j = 2, \dots, N$ in (24).

Theorem 2: For $j = 1, \dots, N$, the point $\delta \mathbf{r}_j^*$ achieves the maximum value in

$$\max_{\delta \mathbf{r}_j \in \mathcal{S}_j} J_j(\boldsymbol{\theta}, \delta \mathbf{r}_j) \quad (25)$$

lies on the boundary of \mathcal{S}_j .

Proof: Since J_j is convex and quadratic in $\delta \mathbf{r}_j$, the proof for each j follows identically to the proof of Theorem 1. We omit the details because of this redundancy. ■

We now consider two special structures for the regions of uncertainty: polyhedral and ellipsoidal regions. For simplicity of exposition, we assume that each region of uncertainty is identical over the possible sensor locations, $\mathcal{S}_1 = \dots = \mathcal{S}_N$. We denote the polyhedral-shaped uncertainty region \mathcal{P} . For the special case of a box-shaped region, we denote the region $\mathcal{B}_{l_x, l_y, l_z}$. This region is parameterized by three values, l_x , l_y ,

and l_z , which indicate the length, width, and height of the region of uncertainty

$$\mathcal{B}_{l_x, l_y, l_z} = \{\delta \mathbf{r} = (\delta x, \delta y, \delta z) : |\delta x| \leq l_x, |\delta y| \leq l_y, |\delta z| \leq l_z\}. \quad (26)$$

For the ellipsoidal region, denoted as $\mathcal{E}_{l_x, l_y, l_z}$, we make use of the same parameters l_x , l_y , and l_z to indicate each of the three axis lengths corresponding to a 3-D ellipsoid. We omit rotation angles from the characterization of the ellipsoid, although, this can be easily incorporated

$$\mathcal{E}_{l_x, l_y, l_z} = \left\{ \delta \mathbf{r} = (\delta x, \delta y, \delta z) : \frac{\delta x^2}{l_x^2} + \frac{\delta y^2}{l_y^2} + \frac{\delta z^2}{l_z^2} \leq 1 \right\}. \quad (27)$$

B. Polyhedral Region of Uncertainty Under $\alpha = 0$

Under $\alpha = 0$, for any region of uncertainty, Theorem 2 allows us to restrict our search in the inner maximization of (24) to the boundary of the region. However, an even stronger statement can be made for the case of polyhedral regions (which include box-shaped regions), as follows.

Theorem 3: Let \mathcal{P} be any polyhedral region of uncertainty. The value $\delta \mathbf{r}_j^*$ achieving the maximum value in

$$\max_{\delta \mathbf{r}_j \in \mathcal{P}} J_j(\boldsymbol{\theta}, \delta \mathbf{r}_j) \quad (28)$$

lies at one of the extreme points (a corner) of \mathcal{P} .

Proof: See Appendix B. ■

As a result of Theorem 3, when $\alpha = 0$, we need only test the extreme points of the polyhedral region of uncertainty in order to determine the overall cost-maximizing perturbation value over the entire region. For the box-shaped region, this implies a search over the eight corner points of the box.

C. Ellipsoidal-Shaped Region of Uncertainty Under $\alpha = 0$

We now turn our attention to the case of an ellipsoidal region of uncertainty. The inner maximization in (24) can be expressed as

$$\max_{\delta \mathbf{r}_j \in \mathcal{E}_{l_x, l_y, l_z}} J_j(\boldsymbol{\theta}, \delta \mathbf{r}_j) = \max_{\delta \mathbf{r}^T \text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\}) \delta \mathbf{r} = 1} \|\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}\|^2 \quad (29)$$

where

$$\text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\}) = \begin{bmatrix} 1/l_x^2 & 0 & 0 \\ 0 & 1/l_y^2 & 0 \\ 0 & 0 & 1/l_z^2 \end{bmatrix}. \quad (30)$$

We will approach this constrained optimization through the Lagrange multiplier method. Using Lagrange multiplier λ , we consider the unconstrained optimization

$$\max_{\delta \mathbf{r}} \left(\|\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}\|^2 + \lambda (\delta \mathbf{r}^T \text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\}) \delta \mathbf{r} - 1) \right). \quad (31)$$

Through a simple matrix derivative, we obtain the optimality condition

$$-2\bar{\mathbf{A}}^T\bar{\mathbf{d}} + 2\left(\lambda \text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\}) + \bar{\mathbf{A}}^T\bar{\mathbf{A}}\right)\delta\mathbf{r} = 0 \quad (32)$$

which implies the optimal value $\delta\mathbf{r}_j^*$ satisfies

$$\delta\mathbf{r}_j^* = \left(\lambda \text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\}) + \bar{\mathbf{A}}^T\bar{\mathbf{A}}\right)^{-1} \bar{\mathbf{A}}^T\bar{\mathbf{d}}. \quad (33)$$

In order to characterize the value of λ such that $\delta\mathbf{r}_j^{*\top}\delta\mathbf{r}_j^* = 1$, simple algebraic manipulation yields the following series of equations:

$$\delta\mathbf{r}_j^{*\top}\text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\})\delta\mathbf{r}_j^* \quad (34)$$

$$= \left\| \text{diag}(\{1/l_x, 1/l_y, 1/l_z\}) \times \left(\lambda \text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\}) + \bar{\mathbf{A}}^T\bar{\mathbf{A}}\right)^{-1} \bar{\mathbf{A}}^T\bar{\mathbf{d}} \right\|^2 \quad (35)$$

$$= \left\| \left(\lambda \mathbf{I} + \hat{\mathbf{A}}^T\hat{\mathbf{A}}\right)^{-1} \hat{\mathbf{A}}^T\bar{\mathbf{d}} \right\|^2 \quad (36)$$

$$= \left\| \mathbf{V}(\lambda \mathbf{I} + \Sigma^T\Sigma)^{-1}\Sigma^T\mathbf{U}\bar{\mathbf{d}} \right\|^2 \quad (37)$$

$$= \bar{\mathbf{d}}^T\mathbf{U}^T\Sigma(\lambda \mathbf{I} + \Sigma^T\Sigma)^{-2}\Sigma^T\mathbf{U}\bar{\mathbf{d}} \quad (38)$$

$$= \frac{\hat{\mathbf{d}}_1^2}{(\lambda + \sigma_1^2)^2} + \frac{\hat{\mathbf{d}}_2^2}{(\lambda + \sigma_2^2)^2} + \frac{\hat{\mathbf{d}}_3^2}{(\lambda + \sigma_3^2)^2}. \quad (39)$$

We have defined $\hat{\mathbf{A}} = \bar{\mathbf{A}}\text{diag}(\{l_x, l_y, l_z\})$. The singular value decomposition of $\hat{\mathbf{A}}$ is given by $\hat{\mathbf{A}} = \mathbf{U}\Sigma\mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are orthogonal matrixes having orthonormal columns. We have additionally defined the 3×1 vector $\hat{\mathbf{d}} = \Sigma^T\mathbf{U}\bar{\mathbf{d}}$. Since we require $\delta\mathbf{r}_j^{*\top}\text{diag}(\{1/l_x^2, 1/l_y^2, 1/l_z^2\})\delta\mathbf{r}_j^* = 1$, (34)–(39) imply that λ must satisfy the sixth-order polynomial equation

$$\hat{\mathbf{d}}_1^2(\lambda + \sigma_2^2)^2(\lambda + \sigma_3^2)^2 + \hat{\mathbf{d}}_2^2(\lambda + \sigma_1^2)^2(\lambda + \sigma_3^2)^2 + \hat{\mathbf{d}}_3^2(\lambda + \sigma_1^2)^2(\lambda + \sigma_2^2)^2 = (\lambda + \sigma_2^2)^2(\lambda + \sigma_1^2)^2(\lambda + \sigma_3^2)^2. \quad (40)$$

Thus, we have reduced the problem in the case of ellipsoidal regions of uncertainty to obtaining the roots of a sixth-order polynomial and testing each of the implied boundary locations for maximum cost in J_j .

D. General Case of $\alpha > 0$: Dynamic-Programming Approach

In this section, we treat the general problem for $\alpha > 0$. A simple dynamic-programming formulation is introduced for the optimal solution to the inner maximization in (18). Whereas the polyhedral and ellipsoidal uncertainty regions under $\alpha = 0$ yield simple and tractable solution approaches to the inner maximization in (18), the general formulation for $\alpha > 0$ can result in a trajectory of uncertainty values $(\delta\mathbf{r}_1, \dots, \delta\mathbf{r}_N)$, where it is not necessarily true that $\delta\mathbf{r}_j$ lies on the boundary of \mathcal{S}_j for $j = 1, \dots, N$.

The dynamic-programming algorithm for achieving maximum cost in the inner maximization of (18) is given by

$$V_1 = \max_{\delta\mathbf{r}_1 \in \mathcal{S}_1} \sum_{k=1}^M \|d_{1,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,1}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,1})\delta\mathbf{r}_1\|^2 + V_2(\delta\mathbf{r}_1) \quad (41)$$

$$V_j(\delta\mathbf{r}_{j-1}) = \max_{\delta\mathbf{r}_j \in \mathcal{S}_j(\delta\mathbf{r}_{j-1})} \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j})\delta\mathbf{r}_j\|^2 + \gamma\|\delta\mathbf{r}_j - \alpha\delta\mathbf{r}_{j-1}\|^2 + V_{j+1}(\delta\mathbf{r}_j), \quad j = 2, \dots, N \quad (42)$$

$$V_{N+1} = 0. \quad (43)$$

Since the uncertainty regions are continuous, we must discretize the set \mathcal{R} in order to execute the DP algorithm. Our approximation will lead to a finite number of possible trajectories. We will demonstrate that our approximation is consistent, in the sense that as our discretization becomes finer, we approach the optimal maximum cost achieved in (41)–(43).

Let \mathbb{Z} be the set of integers. Define the rational lattice \mathcal{L}^n by

$$\mathcal{L}^n = \{\mathbf{r} \in \mathbb{R}^{3N} : n\mathbf{r}_i \in \mathbb{Z}, \quad \text{for } i = 1, \dots, 3N\}. \quad (44)$$

Then, we define the discretized space of uncertainty trajectories

$$\mathcal{R}^n = \mathcal{L}^n \cap \mathcal{R}. \quad (45)$$

This discretization induces regions \mathcal{S}_j^n for $j = 1, \dots, n$, which we define recursively as

$$\mathcal{S}_1^n = \{\delta\mathbf{r}_1 \in \mathbb{R}^3 : (\delta\mathbf{r}_1, \delta\mathbf{r}') \in \mathcal{R}^n\} \quad (46)$$

$$\mathcal{S}_j^n(\delta\mathbf{r}_{j-1}) = \left\{ \delta\mathbf{r}_j \in \mathbb{R}^3 : (\delta\mathbf{r}', \delta\mathbf{r}_{j-1}, \delta\mathbf{r}_j, \delta\mathbf{r}'') \in \mathcal{R}^n, \delta\mathbf{r}' \in \mathbb{R}^{3(j-2)} \right\}, \quad j = 2, \dots, N. \quad (47)$$

Our approximate DP algorithm, parameterized by the discretization index n , is then given for $n = 1, 2, \dots$, by

$$V_1^n = \max_{\delta\mathbf{r}_1 \in \mathcal{S}_1^n} \sum_{k=1}^M \|d_{1,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,1}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,1})\delta\mathbf{r}_1\|^2 + V_2^n(\delta\mathbf{r}_1) \quad (48)$$

$$V_j^n(\delta\mathbf{r}_{j-1}) = \max_{\delta\mathbf{r}_j \in \mathcal{S}_j^n(\delta\mathbf{r}_{j-1})} \sum_{k=1}^M \|d_{j,k} - s_k(\boldsymbol{\theta}, \mathbf{r}_{0,j}) - \mathbf{A}_k(\boldsymbol{\theta}, \mathbf{r}_{0,j})\delta\mathbf{r}_j\|^2 + \gamma\|\delta\mathbf{r}_j - \alpha\delta\mathbf{r}_{j-1}\|^2 + V_{j+1}^n(\delta\mathbf{r}_j), \quad j = 2, \dots, N \quad (49)$$

$$V_{N+1}^n = 0. \quad (50)$$

Theorem 4: The approximation (45) is consistent, in the sense that $V_1^n \rightarrow V_1$ as $n \rightarrow \infty$.

Proof: Suppose the trajectory $\delta\mathbf{r} = (\delta\mathbf{r}_1, \dots, \delta\mathbf{r}_N) \in \mathcal{R}$ achieves the maximum cost in (41)–(43). Let $\varepsilon > 0$ and define

$$\mathcal{B}_\varepsilon = \{\delta\mathbf{r}' \in \mathbb{R}^{3N} : \|\delta\mathbf{r} - \delta\mathbf{r}'\|^2 \leq \varepsilon\}. \quad (51)$$

By the definition of \mathcal{R} , the set $\mathcal{B}_\varepsilon \cap \mathcal{R}$ must be a convex set of nonzero volume. This implies that there must exist a rational-valued vector $\delta\mathbf{r}' \in \mathcal{B}_\varepsilon \cap \mathcal{R}$. Let the lowest common denominator of all the components of $\delta\mathbf{r}'$ be n . Then, clearly, $n\delta\mathbf{r}'$ is integer-valued. Thus, $\delta\mathbf{r}' \in \mathcal{R}^n$. Expressing $\delta\mathbf{r}^n = \delta\mathbf{r} + \Delta\mathbf{r}$, we obtain

$$|J(\boldsymbol{\theta}, \delta\mathbf{r}) - J(\boldsymbol{\theta}, \delta\mathbf{r}^n)| = \left| \|\bar{\mathbf{d}} - \bar{\mathbf{A}}\delta\mathbf{r}\|^2 - \|\bar{\mathbf{d}} - \bar{\mathbf{A}}\delta\mathbf{r}^n\|^2 \right| \quad (52)$$

$$= \left| \|\bar{\mathbf{A}}\Delta\mathbf{r}\|^2 - 2\Delta\mathbf{r}^T \bar{\mathbf{A}}^T (\bar{\mathbf{d}} - \bar{\mathbf{A}}\delta\mathbf{r}) \right| \quad (53)$$

$$\leq \sigma_1^2 \|\Delta\mathbf{r}\|^2 + 2\sigma_1 \|\Delta\mathbf{r}\| \|\bar{\mathbf{d}} - \bar{\mathbf{A}}\delta\mathbf{r}\| \quad (54)$$

$$\leq \sigma_1^2 \varepsilon + 2\sigma_1 \sqrt{\varepsilon} \|\bar{\mathbf{d}} - \bar{\mathbf{A}}\delta\mathbf{r}\| \quad (55)$$

where σ_1 is the maximum singular value of matrix $\bar{\mathbf{A}}$. Noting that the quantity $\|\bar{\mathbf{d}} - \bar{\mathbf{A}}\delta\mathbf{r}\|$ is a constant, we have that the approximation is consistent, as desired. ■

IV. NUMERICAL STUDIES

In this section, we validate the analysis of our previous sections through simulation studies. We consider both the cases of independent and dependent sequences of data-collection locations.

A. Independent Data-Collection Locations: Simulated Data

Our first numerical study makes use of simulated data to test the classification algorithm when there is no knowledge of dependence between adjacent data-collection locations. Thus, the algorithm employs the value $\alpha = 0$. In this scenario, our sensing-system model is comprised of two square colocated transmit and receive coils with dimensions 0.5 m on each side. These coils sample a 1-m square area on an equally spaced 4×4 grid of measurement points. This models the scenario where an object has already been detected and data collection is being conducted in the vicinity of the object to assist in the classification process. Our model assumes that the object to be classified resides at a depth between 0.05 and 2.0 m below the surface of the ground. The sensor is modeled as a frequency-domain sensor, with complex data (in-phase and quadrature) collected at 30 equally logarithmically spaced frequencies between 10 and 30 kHz.

Our simulated library of target objects consists of four objects: a 3 in long \times 1 in diameter stainless-steel cylinder (S1), a 6 in long \times 1 in diameter stainless-steel cylinder (S2), a 3 in long \times 1 in diameter aluminum cylinder (A1), and

a 6 in long \times 1 in diameter aluminum cylinder (A2). The “ground truth” model, for the scattering characteristics of these targets, was obtained using the method of [36], where the dipole model was taken to be exact. Four terms were kept in each λ_i summations in (4), and all expansion coefficients were taken to be one (see [36] for additional details).

Data collection is done with the SNR set at approximately 12 dB, with noise always assumed to be independent identically normally distributed. The SNR for this and all other cases was calculated as follows:

$$\text{SNR} = 10 \log_{10} \frac{\|d\|^2}{\sigma^2 l_d} \quad (56)$$

where d denotes the signal vector, l_d is its length, and σ^2 denotes the noise variance. Throughout this paper, we assume that the noise is independent identically Gaussian distributed. We make use of the “databased” classification method of [17], which uses a classification statistic that essentially tests for a best fit to the object library. The pole characteristics of our simulated targets are the same as those in [17, Sec. IV]. As noted in [17], the pole characteristics vary significantly as a function of object material but much less so between objects of the same material. Consequently, we anticipate significantly better classification of object material than of precise object.

A Monte Carlo approach was used to analyze and compare the performance of our algorithms for this library. Specifically, 100 separate runs were carried out for each object, where we randomized uniformly over object type, object location, orientation, and additive sensor noise (by selecting a new set of parameters according to a uniform distribution at each Monte Carlo run). The corresponding values of the Euler angles were chosen randomly (with a uniform distribution) over their full range of definition (either 0 to 2π or 0 to π depending on the angle [40]).

Our Monte Carlo studies were used to test both the cases of box-shaped and ellipsoidal-shaped uncertainty regions, modeled in (26) and (27), respectively. In the case of box-shaped uncertainty, each uncertainty region was assumed to be 10 cm on each side, while for the ellipse-shaped uncertainty, we assumed a spherical uncertainty region of 10-cm diameter. We compared the classification results against the classification results of the pole-based classification scheme in [17], where the algorithm is agnostic with regard to sensor positional uncertainty.

The classification results for the box-shaped uncertainty regions are summarized in the confusion matrixes of Tables I and II, while those of the ellipsoidal-shaped uncertainty regions are summarized in Tables III and IV. The element on the i th row and j th column of each matrix displays the number of times object i was the true target, and object j was selected by the given processing scheme. Direct comparison between Tables I and II for the box case, and of Tables III and IV for the ellipsoidal case, show dramatic improvements both in correct classification of type and material of the targets under study when positional uncertainties are taken into consideration. For the data generated under box-shaped positional uncertainty, the correct-material-classification rate improves from 25%, 39% (Aluminum targets), 63%, and 58% (Steel targets) to 95%,

TABLE I
CLASSIFICATION RESULTS FOR BOX-SHAPED POSITIONAL
UNCERTAINTY UNDER THE ALGORITHM
THAT IGNORES UNCERTAINTY

	A1	A2	S1	S2
A1	15	9	35	41
A2	17	9	37	37
S1	14	14	47	25
S2	19	17	37	27

TABLE II
CLASSIFICATION RESULTS FOR BOX-SHAPED POSITIONAL
UNCERTAINTY UNDER THE MIN-MAX FORMULATION

	A1	A2	S1	S2
A1	48	39	12	1
A2	36	53	11	0
S1	0	0	12	88
S2	0	0	18	82

TABLE III
CLASSIFICATION RESULTS FOR ELLIPSOIDAL-SHAPED
POSITIONAL UNCERTAINTY UNDER THE ALGORITHM
THAT IGNORES UNCERTAINTY

	A1	A2	S1	S2
A1	12	13	34	41
A2	22	17	28	33
S1	18	19	31	32
S2	20	22	32	25

TABLE IV
CLASSIFICATION RESULTS FOR ELLIPSOIDAL-SHAPED POSITIONAL
UNCERTAINTY UNDER THE MIN-MAX FORMULATION

	A1	A2	S1	S2
A1	66	29	3	2
A2	69	26	3	2
S1	0	0	13	87
S2	0	0	13	87

95%, 100%, and 100%, respectively. Similarly, the correct-material-classification rate in the case of data generated under ellipsoidal-shaped positional uncertainty improves from 24%, 26% (Aluminum targets), 72%, and 64% (Steel targets) to 95%, 95%, 100%, and 100%, respectively.

B. Robustness of the Algorithm

We consider the possibility of having incorrect knowledge of the bounding information for sensor positional uncertainty, in

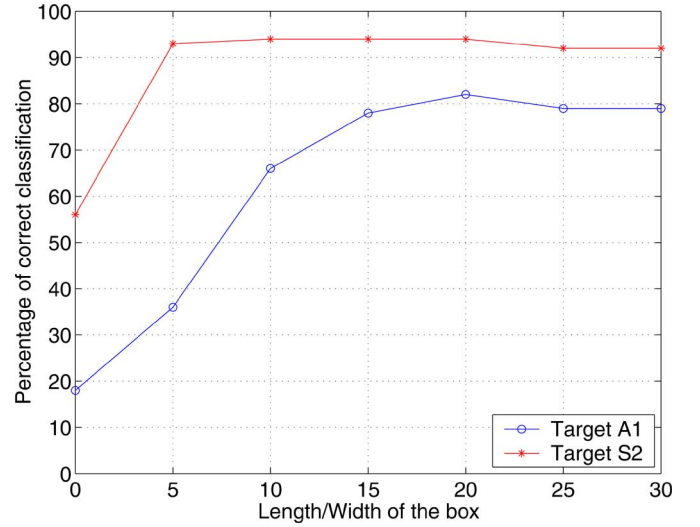


Fig. 3. Plot of percentage of successful classification of object type for varying bounding-box dimensions used within the optimization.

order to determine the robustness of our technique. While this robustness study was only conducted for the case of $\alpha = 0$, we expect similar results under nonzero values of α . We considered objects A1 and S2, detailed in Section IV-A. For each target of interest, we generated data sets in a similar fashion as described in Section IV-A. In these studies, however, we assume that the true error in sensor position is restricted in the x and y directions by 25 cm and in the z direction by 10 cm.

Given each data set, we conducted optimizations assuming a range of bounding boxes. First, we considered a region equal to a point, meaning that we assume no knowledge of any uncertainty in the sensor positions. This assumption is similar to the algorithm presented in [17]. Subsequently, we considered regions with nonzero volume, having z dimension equal to 10 cm in each case, while the x and y dimensions were simultaneously stepped from 5 to 30 cm in increments of 5 cm.

The results of this paper are presented in Fig. 3, where each data point represents the percentage of successful target-type classifications (on the vertical axis) corresponding to a particular bounding-box dimension (on the horizontal axis). Note that the value of 0 cm on the horizontal axis represents the case where our optimization assumed the bounding region to be equal to a point. Each data point is obtained by considering 100 independently generated data sets.

As is clear from Fig. 3, the classifier behaves well for bounding boxes having dimensions in the vicinity of the true bounding box (25 cm in both the x and y directions). We thus conclude that our algorithm is robust to the presence of incorrect bounding information.

C. Dependent Positional Uncertainty: Simulated Data

This simulation study concerns dependent positional uncertainties, where a known or assumed value $\alpha \neq 0$ ties each successive positional uncertainty value to its predecessor. We make use of simulated data. The sensing system and region of interest containing the target to be classified are identical to that in Section IV-A. Data are assumed to be collected over a

TABLE V
CLASSIFICATION RESULTS FOR DEPENDENT POSITIONAL
UNCERTAINTIES FOR THE MIN-MAX FORMULATION
UNDER THE ASSUMPTION OF $\alpha = 0$

	A1	A2	S1	S2
A1	9	15	39	37
A2	13	15	39	33
S1	26	29	20	25
S2	26	26	24	24

TABLE VI
CLASSIFICATION RESULTS FOR DEPENDENT POSITIONAL
UNCERTAINTIES FOR THE APPROXIMATE
DYNAMIC-PROGRAMMING FORMULATION

	A1	A2	S1	S2
A1	59	32	6	3
A2	26	61	8	5
S1	1	6	70	23
S2	0	8	27	65

square grid of 16 data-collection locations, with the sequence of locations following the winding path depicted in Fig. 2. As in Section IV-A, we assume that data are collected over 30 logarithmically spaced frequencies. We make use of the correlation parameter $\alpha = 1$ and assume that each innovation region is a box-shaped region of uncertainty with each edge having 10 cm in length. Recall that the innovation region consists of the set of new positional uncertainty associated with a data-collection location that compounds with the previous location's positional uncertainty.

The simulation results for this paper are provided in the confusion matrixes of Tables V and VI: Table V contains the classification results for the min-max algorithm that assumes $\alpha = 0$, which corresponds to the case of no knowledge of any correlation between adjacent uncertainty values; Table VI contains the classification results for the approximate dynamic program outlined in (48)–(50), with a sampling of the space of positional trajectories at intervals of 5 cm. Clearly, the dynamic-programming approach provides significant performance improvements over the min-max algorithm that operates under the assumption of $\alpha = 0$. Here, the correct-type classification was improved from 51%, 51%, 68%, and 69% for targets A1, A2, S1, and S2, respectively, to 91%, 87%, 93%, and 92%, respectively.

D. Field-Data Studies

Our second target library is comprised of nine target types whose details are shown in Table VII, where we have separated these targets into three sets, designated by letters *L*, *M*, and *S*. The letter *L* corresponds to targets that are deeply buried at a depth ranging from 90 to 110 cm. The letter *M* corresponds to targets buried at a depth ranging from 50 to 80 cm. Finally,

TABLE VII
GEM-3 PIPE-DATA-POLE LIBRARY

Target ID	Description
L1,L2	15.5cm \times 50.8cm Steel Pipe
M1,M2,M5	7.90cm \times 15.7cm Steel Pipe
M3,M4,M7	6.40cm \times 30.5cm Steel Pipe
S1,S10	2.00cm \times 10.2cm Steel Pipe
S3,S5,S6,S8	4.10cm \times 15.2cm Steel Pipe
S4,S7	4.10cm \times 10.2cm Steel Pipe
M6	2.30cm \times 30.5cm Aluminum Pipe
S2,S9	2.30cm \times 15.2cm Aluminum Pipe
S11,S12	2.30cm \times 15.2cm Copper Pipe

the letter *S* corresponds to shallowly buried targets at a depth ranging from 10 to 30 cm. This labeling is consistent with the notation of [17] and [20].

The ground truth and the measured data for this paper were obtained by the authors of [20], where the authors measured the data in a 10 \times 10 m field that contained over 21 metallic targets (listed in Table VII). This test site (detailed in [20]) is located in Raleigh, NC, and was specially designed by Geophex Inc. The data for this sensing system come from the GEM-3 sensor developed by Geophex. This sensor has been used successfully in many environmental sites and can detect small targets, such as UXO and landmines, providing high spatial resolution [20]. The data were obtained using the dead-reckoning method at a line spacing of 25 cm and a height of about 20 cm above the ground. The GEM-3 sensor collected complex-valued frequency-domain data at five different frequencies, in a bandwidth from 30 Hz to 24 kHz. The GEM-3, in this case, collected about 8–10 data points per second, which resulted in a data interval of about 15 cm [20]. According to [20], the positional error for such data could be as large as 20 cm due to uneven walking speed and incorrect walking path. The error associated with sensor height can be up to 5 cm at some points [20]. These values were applied as the respective length, width, and height of the box-shaped positional uncertainty region. Since no notion of the correlation of one uncertainty region to the next was available with the data, we initially proceed with an assumption of $\alpha = 0$.

Classification results for the algorithm presented in [17] demonstrated reasonable performance with regard to this data, even without considering positional uncertainties. However, [17] required distinguishing between deeply and shallowly buried objects and applying appropriate classifiers (respectively, databased and pole-based classifiers). Here, we restrict our min-max formulation to databased classification and compare against the algorithm of [17], when exclusively databased classification is employed.

The classification results for the field data are shown in Table VIII. It is clear from the table that our min-max formulation is significantly better than the algorithm of [17] at correctly classifying the target objects. Overall, we see 47.62% of all targets correctly classified under the algorithm of [17], while our min-max formulation achieves 71.43% correct classification.

TABLE VIII
CLASSIFICATION RESULTS FOR GEM-3 DATA. HERE, “C” DENOTES
A CORRECTLY CLASSIFIED TARGET, WHILE “M” DENOTES
A MISCLASSIFIED TARGET

Actual Target	Algorithm 1:		Algorithm 2:	
	Ignores uncertainty		Min-max Box formulation	
	Estimated	Result	Estimated	Result
L1	L1L2	C	L1L2	C
L2	L1L2	C	L1L2	C
M1	M1M2M5	C	M1M2M5	C
M2	M1M2M5	C	M1M2M5	C
M3	M3M4M7	C	M3M4M7	C
M4	M3M4M7	C	M3M4M7	C
M5	M1M2M5	C	M1M2M5	C
M6	M6	C	M6	C
M7	M6	M	M3M4M7	C
S1	S3S5S6S8	M	S4S7	M
S2	M3M4M7	M	S1S10	M
S3	M1M2M5	M	S3S5S6S8	C
S4	S4S7	C	S4S7	C
S5	M3M4M7	M	S3S5S6S8	C
S6	M6	M	S11S12	M
S7	S4S7	C	S4S7	C
S8	S4S7	M	S3S5S6S8	C
S9	S4S7	M	M6	M
S10	M6	M	L1L2	M
S11	M6	M	S11S12	C
S12	S1S10	M	S3S5S6S8	M

Since the field data collected in [20] were obtained using the dead-reckoning approach, one would expect that the successive sensor positions would incur accumulated positional uncertainty. Consequently, a nonzero value of α could be used within a dynamic-programming optimization approach. Unfortunately, the data collected in [20] only have a fixed box-shaped uncertainty region associated with each data-collection location, which fits well with having parameter $\alpha = 0$. The limited information regarding the dependence of uncertainty from one location to the next implies that we can, at best, guess the structure of the successive innovation regions from one sensor position to the next. Nevertheless, in order to provide an additional test of our dynamic-programming approach, we have tested several objects listed in Table VIII, assuming that the sensor positional uncertainty from one location to the next can be perturbed by no more than 5 cm. We considered objects L1, M1, M6, S1, and S4 within this dynamic program framework and obtained identical classification results to those of the box algorithm (presented in Table VIII).

V. CONCLUSION

In this paper, we introduced a model-based algorithm that explicitly accounts for the positional uncertainties involved

with data collection using EMI sensors. We utilized a forward model for the EMI sensing system that relates a small set of parameters to the observed data. Although we present this paper in the context of classifying buried objects under the assumption of EMI sensor positional uncertainty, our model applies quite generally to other sensing modalities, as well as to other inverse optimization goals.

Assuming that data are collected at locations having positional uncertainty, we approached the problem by looking for model parameters that minimize the maximum mismatch to the data. This maximum was taken over all possible sensor locations, which we assumed to be bounded within some finite region of the space. For the case of independent uncertainty regions at successive data-collection locations, we analyzed general polyhedral regions and ellipsoidal regions of uncertainty and determined simple characterizations of the min-max formulation over these regions. For the case of dependent uncertainty regions, we introduced a dynamic-programming formulation for determining the optimized min-max classifier output.

Our numerical studies based on simulated data point to significant benefits in applying our min-max strategies. Classification of true field data also yielded improvements over approaches that do not take positional uncertainty into account. An important avenue for future research lies in addressing the heavy computational burden of applying dynamic programming for the case of dependent positional uncertainty. There is a body of literature for approximate dynamic programming that would be effective to this end (see [48] for an introduction to approximate dynamic-programming techniques). We defer discussion of such techniques from this paper as our contribution has been to introduce dynamic programming as a useful tool for addressing the problem of positional uncertainty in our classification problem. Additionally, an important avenue for future work is in the application of our model for dependent positional uncertainty to field data.

While this paper has focused on the issue of uncertain sensor positions, other forms of uncertainty also arise in the classification process. In particular, it is well understood in the classification problem we have considered here that the decay rates/resonances also exhibit uncertainties [45], [46]. This property is most relevant in the design of classifiers that process poles after they have been estimated from the data and can be helpful in designing classifiers that are more robust. More recently, Miller *et al.* in [45] and [46] have explored how the pole clouds can be used basically as regularizers in the estimation process. Since we have focused in this paper on developing mechanism for addressing the sensor-positional-uncertainty problem in general classification settings, we have left this particular application as an interesting area of future research. Another avenue for future research includes the extension of the classifier algorithm such that it will be able to identify clutter items as objects that are not in the library. This step is important in real-world UXO and demining problems, where there is a strong desire to correctly reject clutter items. Finally, as discussed in Section III, the automatic determination of the parameter γ remains as an avenue for future research.

APPENDIX A PROOF OF THEOREM 1

Suppose that $\delta \mathbf{r}^* = (\delta \mathbf{r}_1^*, \dots, \delta \mathbf{r}_N^*)$ is interior to \mathcal{R} , which implies that there exists $\epsilon > 0$ such that

$$\left\{ (\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) : \sum_{i=1}^N |\delta \mathbf{r}_i - \delta \mathbf{r}_i^*| \leq \epsilon \right\} \subseteq \mathcal{R}. \quad (57)$$

We will demonstrate that there exists a point on the boundary of \mathcal{R} having cost greater or equal to that achieved at $\delta \mathbf{r}^*$.

Expanding (17), we have

$$J(\boldsymbol{\theta}, \delta \mathbf{r}_1, \dots, \delta \mathbf{r}_N) = \bar{\mathbf{d}}^T \bar{\mathbf{d}} - 2\delta \mathbf{r}^T \bar{\mathbf{A}}^T \bar{\mathbf{d}} + \delta \mathbf{r}^T \bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r}. \quad (58)$$

The derivative of J with respect to $\delta \mathbf{r}$ is then given by

$$\frac{dJ}{d\delta \mathbf{r}} = -2\bar{\mathbf{A}}^T \bar{\mathbf{d}} + 2\bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r} \quad (59)$$

which implies that the direction of the maximum cost improvement at the point $\delta \mathbf{r}^*$ is oriented toward vector

$$-2\bar{\mathbf{A}}^T \bar{\mathbf{d}} + 2\bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r}^*. \quad (60)$$

Since $\delta \mathbf{r}^*$ is interior to \mathcal{R} and because \mathcal{R} is assumed to be compact, there must exist a positive real number κ such that for any $\epsilon \in (0, \kappa]$

$$\delta \mathbf{r}^* + \epsilon (\bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r}^* - \bar{\mathbf{A}}^T \bar{\mathbf{d}}) \in \mathcal{R} \quad (61)$$

and such that there exists $\epsilon' > 0$ and such that for $\epsilon \in (0, \epsilon']$

$$\delta \mathbf{r}^* + (\kappa + \epsilon) (\bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r}^* - \bar{\mathbf{A}}^T \bar{\mathbf{d}}) \notin \mathcal{R}. \quad (62)$$

We consider the boundary point of \mathcal{R} , given by

$$\delta \mathbf{r}' = \delta \mathbf{r}^* + \kappa (\bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r}^* - \bar{\mathbf{A}}^T \bar{\mathbf{d}}). \quad (63)$$

The cost at $\delta \mathbf{r}'$ satisfies the following series of relations:

$$J(\boldsymbol{\theta}, \delta \mathbf{r}) = \left\| \bar{\mathbf{d}} - \bar{\mathbf{A}} (\delta \mathbf{r}^* + \kappa \bar{\mathbf{A}}^T \bar{\mathbf{A}} \delta \mathbf{r}^* - \kappa \bar{\mathbf{A}}^T \bar{\mathbf{d}}) \right\|^2 \quad (64)$$

$$= \left\| \left(\mathbf{I} + \kappa \bar{\mathbf{A}} \bar{\mathbf{A}}^T \right) (\bar{\mathbf{d}} - \bar{\mathbf{A}} \delta \mathbf{r}^*) \right\|^2 \quad (65)$$

$$= (\bar{\mathbf{d}} - \bar{\mathbf{A}} \delta \mathbf{r}^*)^T \left(\mathbf{I} + 2\kappa \bar{\mathbf{A}} \bar{\mathbf{A}}^T + \kappa^2 \bar{\mathbf{A}} \bar{\mathbf{A}}^T \bar{\mathbf{A}} \bar{\mathbf{A}}^T \right) (\bar{\mathbf{d}} - \bar{\mathbf{A}} \delta \mathbf{r}^*) \quad (66)$$

$$\geq (\bar{\mathbf{d}} - \bar{\mathbf{A}} \delta \mathbf{r}^*)^T (\bar{\mathbf{d}} - \bar{\mathbf{A}} \delta \mathbf{r}^*) \quad (67)$$

$$= J(\boldsymbol{\theta}, \delta \mathbf{r}^*). \quad (68)$$

The inequality relation in (67) follows from (66), because the matrixes $\kappa \bar{\mathbf{A}} \bar{\mathbf{A}}^T$ and $\bar{\mathbf{A}} \bar{\mathbf{A}}^T \bar{\mathbf{A}} \bar{\mathbf{A}}^T$ must each be positive semidefinite. Thus, these matrixes can only contribute positive cost to $J(\boldsymbol{\theta}, \delta \mathbf{r}')$. From (68), we have that the boundary point $\delta \mathbf{r}'$ of \mathcal{R} achieves greater or equal cost to the interior point $\delta \mathbf{r}^*$.

APPENDIX B PROOF OF THEOREM 3

Using Theorem 2, we may, without loss of generality, assume that $\delta \mathbf{r}_j^*$ lies on the boundary of \mathcal{P} . The proof proceeds in two parts: First, we demonstrate that if $\delta \mathbf{r}_j^*$ is interior to a face of \mathcal{P} , then there exists a point along one of the edges of the face at which the cost is greater or equal to that at $\delta \mathbf{r}_j^*$. Subsequently, we demonstrate that if $\delta \mathbf{r}_j^*$ lies on the edge of one of the faces of \mathcal{P} , then there exists an extreme point at which the cost is greater or equal to that at $\delta \mathbf{r}_j^*$.

Assume that $\delta \mathbf{r}_j^*$ is interior to one of the faces of \mathcal{P} . By the structure of J_j in (21), the direction of cost improvement at location $\delta \mathbf{r}_j^*$ is given [as in (60)] by

$$\bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* - \bar{\mathbf{A}}_j^T \bar{\mathbf{d}}_j. \quad (69)$$

Let \mathbf{e}_1 and \mathbf{e}_2 be orthonormal vectors lying in the plane containing the face of \mathcal{P} on which $\delta \mathbf{r}_j^*$ lies. The direction of cost improvement along the plane spanned by \mathbf{e}_1 and \mathbf{e}_2 is obtained through a simple projection and equals

$$\delta \mathbf{r} = \mathbf{e}_1^T (\bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* - \bar{\mathbf{A}}_j^T \bar{\mathbf{d}}_j) \mathbf{e}_1 + \mathbf{e}_2^T (\bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* - \bar{\mathbf{A}}_j^T \bar{\mathbf{d}}_j) \mathbf{e}_2. \quad (70)$$

We must consider two cases. First, if $\delta \mathbf{r} = 0$, then let $\kappa > 0$ be the maximum real number such that $\delta \mathbf{r}^* + \kappa \mathbf{e}_1 \in \mathcal{P}$. Then

$$\begin{aligned} J_j(\boldsymbol{\theta}, \delta \mathbf{r}_j^* + \kappa \mathbf{e}_1) &= \left\| \bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* \right\|^2 - 2\kappa \mathbf{e}_1^T \\ &\quad \times (\bar{\mathbf{A}}_j^T \bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \delta \mathbf{r}_j^*) + \kappa^2 \mathbf{e}_1^T \bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \mathbf{e}_1 \end{aligned} \quad (71)$$

$$= \left\| \bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* \right\|^2 + \kappa^2 \mathbf{e}_1^T \bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \mathbf{e}_1 \quad (72)$$

$$\geq \left\| \bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* \right\|^2 \quad (73)$$

$$= J_j(\boldsymbol{\theta}, \delta \mathbf{r}^*). \quad (74)$$

Above, (72) follows since $\delta \mathbf{r} = 0$, which implies \mathbf{e}_1 is orthogonal to $(\bar{\mathbf{A}}_j^T \bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j^T \bar{\mathbf{A}}_j \delta \mathbf{r}_j^*)$. Thus, (74) implies that there exists a point on the edge of the face achieving equal or better cost than at $\delta \mathbf{r}^*$. For the second case, suppose that $\Delta \mathbf{r} \neq 0$. In this case, let $\kappa > 0$ be the maximum real number such that $\delta \mathbf{r}^* + \kappa \delta \mathbf{r} \in \mathcal{P}$. Then

$$\begin{aligned} J_j(\boldsymbol{\theta}, \delta \mathbf{r}^* + \kappa \delta \mathbf{r}) &= \left\| \left(\mathbf{I} + 2\kappa \bar{\mathbf{A}}_j \mathbf{e}_1 \mathbf{e}_1^T \bar{\mathbf{A}}_j^T + 2\kappa \bar{\mathbf{A}}_j \mathbf{e}_2 \mathbf{e}_2^T \bar{\mathbf{A}}_j^T \right) \right. \\ &\quad \times (\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^*) \left. \right\|^2 \end{aligned} \quad (75)$$

$$= \left\| (\mathbf{I} + \mathbf{Q}) (\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^*) \right\|^2 \quad (76)$$

$$\begin{aligned} &= (\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^*)^T (\mathbf{I} + 2\mathbf{Q} + \mathbf{Q}\mathbf{Q}) \\ &\quad \times (\bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^*) \end{aligned} \quad (77)$$

$$\geq \left\| \bar{\mathbf{d}}_j - \bar{\mathbf{A}}_j \delta \mathbf{r}_j^* \right\|^2 \quad (78)$$

$$= J_j(\boldsymbol{\theta}, \delta \mathbf{r}^*). \quad (79)$$

Above, (76) introduces the symmetric positive semidefinite matrix $\mathbf{Q} = 2\kappa \bar{\mathbf{A}}_j \mathbf{e}_1 \mathbf{e}_1^T \bar{\mathbf{A}}_j^T + 2\kappa \bar{\mathbf{A}}_j \mathbf{e}_2 \mathbf{e}_2^T \bar{\mathbf{A}}_j^T$. Then, (76) follows because \mathbf{Q} and $\mathbf{Q}\mathbf{Q}$ are positive semidefinite. From (78), we have that the cost at point $\delta\mathbf{r}^* + \kappa\delta\mathbf{r}$, which resides on the edge of the face, equal or greater to that at $\delta\mathbf{r}^*$. Thus, we can now conclude that there is always a point on the edge of the face in which \mathbf{r}^* resides, at which the cost is equal or greater to that at \mathbf{r}^* .

The second part of the proof deals with $\delta\mathbf{r}^*$ on an edge of a face of \mathcal{P} . In this case, we assume that vector \mathbf{e}_3 is parallel to the line spanning the edge. Proceeding in a manner similar to the above analysis for a face, we can show that there is an extreme point at which the cost is greater or equal to that at $\delta\mathbf{r}^*$. We omit the details of this part of the proof because of its redundancy.

ACKNOWLEDGMENT

The authors would like to thank L. Carin (Duke University), as well as I. J. Won and H. Huang (Geophex Ltd.) for providing us with the data used in this paper. They would also like to thank the reviewers for their comments and suggestions, which have let to significant improvement in the quality of this paper.

REFERENCES

- [1] B. Barrow and H. H. Nelson, "Model-based characterization of electromagnetic induction signatures obtained with the MTADS electromagnetic array," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1279–1285, Jun. 2001.
- [2] T. H. Bell, B. J. Barrow, and J. T. Miller, "Subsurface discrimination using electromagnetic induction sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1286–1293, Jun. 2001.
- [3] L. H. Carin, Y. Yu, A. R. Dalichaouch, A. R. Perry, P. V. Czipott, and C. Baum, "On the wideband EMI response of rotationally symmetric permeable and conducting target," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1206–1213, Jun. 2001.
- [4] R. Chesney, Y. Das, J. E. McFee, and M. Ito, "Identification of metallic spheroids by classification of their electromagnetic responses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 809–820, 1984.
- [5] L. Collins, P. Gao, and L. Carin, "An improved Bayesian decision theoretic approach for landmine detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1352–1361, 2000.
- [6] S. J. Norton and I. J. Won, "Identification of buried unexploded ordnance from broadband electromagnetic induction data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2253–2261, Oct. 2001.
- [7] T. H. Bell, B. Barrow, and N. Khadr, "Shape-based classification and discrimination of subsurface objects using electromagnetic induction," in *Proc. Int. Geosci. Remote Sens. Symp.*, 1998, pp. 509–513.
- [8] L. S. Riggs, J. E. Mooney, and D. E. Lawrence, "Identification of metallic mine-like objects using low frequency magnetic fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 56–66, Jan. 2001.
- [9] S. L. Tantom and L. M. Collins, "A comparison of algorithms for subsurface target detection and identification using time domain electromagnetic induction data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1299–1306, Jun. 2001.
- [10] H. Nelson and J. R. McDonald, "Multisensor towed array detection system for UXO detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1139–1145, Jun. 2001.
- [11] K. Sun, K. O'Neill, F. Shubitidze, I. Shamatava, and K. D. Paulsen, "Fast data-derived fundamental spheroidal excitation models with application to UXO discrimination," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2573–2583, Nov. 2005.
- [12] L. M. Collins, Y. Zhang, J. Li, H. Wang, L. Carin, S. Hart, S. Rose-Pehrsson, H. Nelson, and J. McDonald, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 347–359, Jan. 1999.
- [13] S. L. Tantom, Y. Q. Wang, and L. M. Collins, "Statistical and adaptive signal processing for UXO discrimination for next-generation sensor data," in *Proc. Prog. Electromagn. Res. Symp.*, 2006, p. 362.
- [14] S. L. Tantom, Y. Yu, Q. Zhu, Y. Q. Wang, and L. M. Collins, "Mitigating measurement position uncertainty to improve UXO detection," in *Proc. UXO Forum*, 2006.
- [15] Y. Yu, S. L. Tantom, and L. M. Collins, "Bayesian mitigation of sensor position errors to improve unexploded ordnance detection," *IEEE Trans. Geosci. Remote Sens.*, submitted for publication June 19, 2006 to IEEE Geoscience and Remote Sensing Letters.
- [16] J. T. Smith and H. F. Morrison, "Optimizing receiver configurations for resolution of equivalent dipole polarizabilities *in situ*," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 7, pp. 1490–1498, Jul. 2005.
- [17] A. B. Tarokh, E. L. Miller, I. J. Won, and H. Huang, "Statistical classification of buried objects from spatially sampled time or frequency domain EMI data," *Radio Sci.*, vol. 39, no. 4, pp. RS4S05.1–RS4S05.11, 2004.
- [18] E. L. Miller, "On some options for statistical classification of buried objects from spatially sampled time or frequency domain EMI data," in *Proc. SPIE—Conf. Detect. Mines and Mine Like Targets: VI*, Apr. 2001, pp. 97–107.
- [19] D. Butler, "Report on a workshop on electromagnetic induction methods for UXO detection and discrimination," *Lead. Edge*, vol. 23, no. 8, pp. 766–770, Aug. 2004.
- [20] H. Huang and I. J. Won, "Characterization of UXO-like targets using broadband electromagnetic induction sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 652–663, Mar. 2003.
- [21] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, "Parameter estimation in the presence of bounded modeling errors," *IEEE Signal Process. Lett.*, vol. 4, no. 7, pp. 195–197, Jul. 1997.
- [22] J. Garnett, *Bounded Analytic Functions*. New York: Springer-Verlag, 2006.
- [23] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 6, pp. 883–893, 1980.
- [24] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for non-orthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, Feb. 1970.
- [25] B. Hassibi, A. H. Sayed, and T. Kailath, "Recursive linear estimation in Krein spaces—Part I: Theory," *IEEE Trans. Autom. Control*, vol. 41, no. 1, pp. 18–33, Jan. 1996.
- [26] —, *Indefinite Quadratic Estimation and Control: A Unified Approach to H_2 and H^∞ Theories*. Philadelphia, PA: SIAM, 1999.
- [27] S. V. Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*. Philadelphia, PA: SIAM, 1991.
- [28] P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an H^∞ -setting," *IEEE Trans. Autom. Control*, vol. 36, no. 2, pp. 152–166, Feb. 1991.
- [29] C. L. Lawson and R. J. Hanson, *Solving Least-Squares Problems*. Philadelphia, PA: SIAM, 1995.
- [30] A. H. Sayed and S. Chandrasekaran, "Parameter estimation with multiple sources and levels of uncertainties," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 680–692, Mar. 2000.
- [31] A. H. Sayed, B. Hassibi, and T. Kailath, "Fundamental inertia conditions for minimization of quadratic forms in indefinite metric spaces," in *Oper. Theory Adv. Appl.*, vol. 87. Basel, Switzerland: Birkhauser, 1996, pp. 309–347.
- [32] U. Shaked and Y. Theodor, " H^∞ -optimal estimation: A tutorial," in *Proc. IEEE Conf. Decis. and Control*, Dec. 1992, pp. 2278–2286.
- [33] M. Byrtek, F. O'Sullivan, M. Muzi, and A. Spence, "An adaptation of ridge regression for improved estimation of kinetic model parameters from pet studies," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 1, pp. 63–68, Feb. 2005.
- [34] R. J. Kelly, "Reducing geometric dilution of precision using ridge regression," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 1, pp. 154–168, Jan. 1990.
- [35] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley-Interscience, 2001.
- [36] P. Gao, L. Collins, P. M. Garber, N. Geng, and L. Carin, "Classification of landmine-like metal targets using wideband electromagnetic induction," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1352–1361, May 2000.
- [37] P. Gao and L. Collins, "A comparison of optimal and suboptimal processors for classification of buried metal objects," *IEEE Signal Process. Lett.*, vol. 6, no. 8, pp. 216–218, Aug. 1999.
- [38] —, "A theoretical performance analysis and simulation of time-domain EMI sensor data for land mine detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 4, pp. 2042–2055, Jul. 2000.

- [39] Y. Das, J. E. McFee, J. Toews, and G. C. Stuart, "Analysis of an electromagnetic induction detector for real-time localization of buried objects," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 3, pp. 278–287, May 1990.
- [40] S. Hassani, *Foundations of Mathematical Physics*. Boston, MA: Allyn and Bacon, 1991.
- [41] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 347–359, Jan. 1999.
- [42] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*. Reading, MA: Addison-Wesley, 1994.
- [43] E. L. Miller and W. C. Karl, *Fundamentals of Inverse Problems*, Fall 2003. Northeastern University course ECEG398 Lecture Notes.
- [44] P. C. Hansen, "The L-curve and its use in the numerical treatment of inverse problems," in *Computational Inverse Problems in Electrocardiology*, ser. Advances in Computational Bioengineering. Southampton, U.K.: WIT Press, 2001, pp. 119–142. 5.
- [45] J. Stalnaker, A. Aliamiri, and E. L. Miller, "An enhanced dipole model for UXO discrimination," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2006.
- [46] A. Aliamiri, J. Stalnaker, and E. L. Miller, "A Bayesian approach for classification of buried objects using non-parametric prior model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2006, pp. 3911–3915.
- [47] G. Eting, *Salas and Hille's Calculus*. Hoboken, NJ: Wiley, 1995.
- [48] D. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 2005.



Eric L. Miller (S'90–M'95–SM'03) received the S.B., S.M., and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1990, 1992, and 1994, respectively, all in electrical engineering and computer science.

He is currently a Professor with the Department of Electrical and Computer Engineering, Tufts University, Medford, MA. His research interests include physics-based tomographic-image formation and object characterization, inverse problems, in general, and inverse scattering, in particular, regularization,

statistical signal and imaging processing, and computational physical modeling. This work has been carried out in the context of applications including medical imaging, nondestructive evaluation, environmental monitoring and remediation, landmine and unexploded-ordnance remediation, and automatic target detection and classification.

Dr. Miller is the recipient of the CAREER Award from the National Science Foundation, in 1996, and the Outstanding Research Award from the College of Engineering at Northeastern University, in 2002. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and was in the same position at the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1998 to 2002. He is the Cogeneral Chair of the 2008 IEEE International Geoscience and Remote Sensing Symposium to be held in Boston, MA. He is a member of Tau Beta Pi, Phi Beta Kappa, and Eta Kappa Nu.



Ashley B. Tarokh (S'00–M'06) received the B.S. degree in electrical engineering from Fairleigh Dickinson University, Teaneck-Hackensack, NJ, in 2000 and the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Boston, MA, in 2003 and 2005, respectively.

She is currently a Postdoctoral Research Fellow in the Department of Radiology, Brigham and Women's Hospital, Harvard Medical School. Her research is focused on developing advanced signal-processing techniques for biomedical-imaging applications. Her

research interests include biomedical signal processing, bioinformatics, signal/image processing (particularly, geophysical signal processing and tomographic imaging), and digital-communication theory.