



(51) International Patent Classification:

G06F 17/10 (2006.01) G01V 3/38 (2006.01)
G01R 27/00 (2006.01) G06F 17/11 (2006.01)
G01V 1/30 (2006.01)

#202 - 2705 Osoyoos Crescent, Vancouver, British Columbia V6T 1X7 (CA).

(21) International Application Number:

PCT/CA2011/000707

(22) International Filing Date:

15 June 2011 (15.06.2011)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/356,918 21 June 2010 (21.06.2010) US

(71) Applicant (for all designated States except US): **UNIVERSITY OF BRITISH COLUMBIA** [CA/CA]; University-Industry Liason Office, #103 - 6190 Agronomy Road, Vancouver, British Columbia V6T 1Z3 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HERMANN, Felix** [CA/CA]; 2735 W. 5th Avenue, Vancouver, British Columbia V6K 1T5 (CA). **HABER, Eldad** [CA/CA];

(74) Agents: **LEE, Brian, Y.** et al.; Gowling Lafleur Henderson LLP, Suite 2300, 550 Burrard Street, Vancouver, British Columbia V6C 2B5 (CA).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR DETERMINING AN ESTIMATED MODEL PARAMETER FROM A MEASURED FIELD

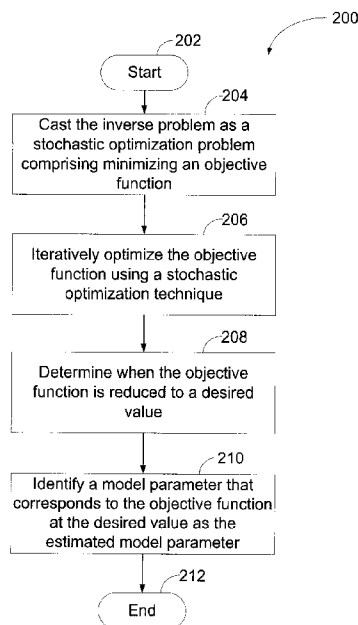


FIG. 2

(57) Abstract: The present disclosure is directed at a method and system for determining an estimated model parameter from a measured field. The measured field may be generated using, for example, multiple sources located through a domain of interest, such as a certain volume of land, and measured using, for example, multiple receivers also located within the domain of interest. The method includes casting an inverse problem relating the measured field to the model parameter as a stochastic optimization problem, in which the stochastic optimization problem includes minimizing an objective function having a misfit expressed as an expected value of a function that depends on a random variable; iteratively optimizing the objective function using a stochastic optimization technique; determining when the objective function is reduced to a desired value; and when the objective function is reduced to the desired value, identifying a model parameter that corresponds to the objective function at the desired value as the estimated model parameter. Exemplary model parameters that can be estimated include conductivity and seismic velocity. Beneficially, using stochastic as opposed to deterministic techniques can result in savings in computational resources in the forms of FLOPS and memory.



LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, **Published:**

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, — *with international search report (Art. 21(3))*
GW, ML, MR, NE, SN, TD, TG).

METHOD AND SYSTEM FOR DETERMINING AN ESTIMATED MODEL PARAMETER FROM A MEASURED FIELD

TECHNICAL FIELD

[0001] The present disclosure is directed at a method and system for determining an estimated model parameter, such as a conductivity model, from a measured field, such as a static electric field. More particularly, the present disclosure is directed at a method and system for determining an estimated model parameter by employing stochastic optimization instead of deterministic optimization, resulting in a technical benefit of savings in computational resources.

BACKGROUND

[0002] Solving parameter estimation problems that employ partial differential equations as constraints and that have multiple right hand sides is performed relatively commonly in applications such as those related to impedance tomography, DC resistivity, electromagnetic imaging, seismic imaging, and hydrology. Such problems are typically mathematically expressed as follows:

$$\begin{aligned} \min \quad & J(y_1, \dots, y_n, u) = \frac{1}{2N} \sum_j \|Py_j - d_j\|^2 + R(u) \\ \text{s.t} \quad & c_j(y_j, u) = A(u)y_j - q_j = 0 \quad j = 1, \dots, N \end{aligned} \quad (1)$$

where $u \in \mathbf{R}^m$ is the discretized model (or control) and $y_j \in \mathbf{R}^n$ $j=1, \dots, N$ is the discretized field (or state). The constraints $c_j(y_j, u) = 0$ are discretized partial differential equations that share the same model u and are linear with respect to q_j . $A(u)$ is presumed to be invertible for all relevant u . The objective function J is composed of two functions: the first is the misfit that measures the fit of y_j , projected by the matrix P to some given data d_j ; and the second, $R(u)$, is assumed to be a convex regularization functional. Every row in P represents a measurement at a certain point in space and time.

[0003] Solving Equation (1) conventionally is very computationally intensive in terms of both FLOPS and memory. Accordingly, research and development continues into methods and

apparatuses for solving such parameter estimation problems in more computationally efficient ways.

SUMMARY

[0004] According to a first aspect, there is provided a method for determining an estimated model parameter from a measured field. The method includes casting an inverse problem relating the measured field to the model parameter as a stochastic optimization problem, wherein the stochastic optimization problem comprises minimizing an objective function comprising a misfit expressed as an expected value of a function that depends on a random variable; iteratively optimizing the objective function using a stochastic optimization technique; determining when the objective function is reduced to a desired value; and when the objective function is reduced to the desired value, identifying a model parameter that corresponds to the objective function at the desired value as the estimated model parameter.

[0005] The measured field may intersect a medium. The method may accordingly also include reconstructing an image of the medium from the estimated model parameter.

[0006] The method may also include generating the measured field using a source, and measuring the measured field using a receiver. Different sources and receivers may be used in different aspects. For example, to perform seismic tomography, the sources may be placed in boreholes and generate a pressure field or a displacement field in response to an acoustic signal, the receivers may be placed on the surface, and the estimated model parameter may be seismic velocity. In another aspect and potential application, the sources may generate a magnetic field in response to a current signal and the estimated model parameter may be conductivity. In another aspect and potential application, the sources may generate a static electric field in response to a current signal, and the estimated model parameter may be conductivity. Different applications are possible in a variety of aspects so long as estimating the model parameter from the measured field is an inverse problem that can be cast as a stochastic optimization problem according to the aspects described herein.

[0007] At least one of the receivers may not measure data from at least one of the sources. This may occur because of relative motion between the sources and receivers.

[0008] The stochastic optimization technique applied to the stochastic optimization problem may be sample average approximation, and alternatively may be stochastic approximation.

[0009] According to another aspect, there is provided a system for determining an estimated model parameter from a measured field. The system includes a processor; and a memory communicatively coupled to the processor and having encoded thereon statements and instructions to cause the processor to execute a method, which includes casting an inverse problem relating the measured field to the model parameter as a stochastic optimization problem, wherein the stochastic optimization problem comprises minimizing an objective function comprising a misfit expressed as an expected value of a function that depends on a random variable; iteratively optimizing the objective function using a stochastic optimization technique; determining when the objective function is reduced to a desired value; and when the objective function is reduced to the desired value, identifying a model parameter that corresponds to the objective function at the desired value as the estimated model parameter.

[0010] The system may also include a display communicatively coupled to the processor, and the method may also include reconstructing an image of a medium that the measured field intersects from the estimated model parameter.

[0011] The system can also include one or more sources, communicatively coupled to the processor and located at source positions, configured to emit the measured field when stimulated using a signal; and one or more receivers, communicatively coupled to the processor and located at receiver positions separated from the source positions, configured to measure the measured field. Different types of sources and receivers can be used. For example, in one aspect and to perform seismic tomography, the sources may be placed in boreholes and configured to generate a pressure field or a displacement field in response to an acoustic signal, the receivers may be placed on the surface, and the estimated model parameter may be seismic velocity. In another aspect and potential application, the sources may be configured to generate a magnetic field in response to a current signal and the estimated model parameter may be conductivity. In another aspect and potential application, the sources may be configured to generate a static electric field in response to a current signal, and the estimated model parameter may be conductivity.

Different applications are possible in a variety of aspects so long as estimating the model parameter from the measured field is an inverse problem that can be cast as a stochastic optimization problem according to the aspects described herein.

[0012] At least one of the receivers may not measure data from at least one of the sources. This may be due, for example, to relative motion between the sources and receivers.

[0013] The stochastic optimization technique applied to the stochastic optimization problem may be sample average approximation, and alternatively may be stochastic approximation.

[0014] According to another aspect, there is provided a computer readable medium having encoded thereon statements and instructions to cause a processor to execute a method of any of the foregoing aspects or any useful combination thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] In the accompanying drawings, which illustrate one or more exemplary embodiments:

[0016] Figure 1 depicts a system for determining an estimated model parameter from a measured field, according to one embodiment.

[0017] Figure 2 depicts a method for determining an estimated model parameter from a measured field, according to another embodiment.

[0018] Figures 3(a) – (d) depict approximations to the misfit using different numbers of vectors, which can form part of the embodiments of Figures 1 and 2.

[0019] Figure 4 depicts a schematic of a system used to perform electromagnetic impedance tomography, according to another embodiment.

[0020] Figures 5(a) – (d) depicts results of the electromagnetic impedance tomography performed using the system of Figure 4.

[0021] Figure 6 depicts a seismic tomography experiment in which a model parameter to be estimated is seismic velocity, and which shows both sources and receivers.

[0022] Figures 7(a) – (f) depict results of the seismic tomography experiment of Figure 6.

DETAILED DESCRIPTION

[0023] Directional terms such as “top,” “bottom,” “upwards,” “downwards,” “vertically” and “laterally” are used in the following description for the purpose of providing relative reference only, and are not intended to suggest any limitations on how any article is to be positioned during use, or to be mounted in an assembly or relative to an environment.

[0024] In several fields, such as impedance tomography, DC resistivity, electromagnetic imaging, seismic imaging, and hydrology, tangible phenomena can be imaged by performing parameter estimation. The parameters that are estimated are those of a model, such as a conductivity model, for a particular domain of interest. Such parameter estimation can be expressed mathematically as Equation (1), which is reproduced below:

$$\begin{aligned} \min \quad & J(y_1, \dots, y_n, u) = \frac{1}{2N} \sum_j \|Py_j - d_j\|^2 + R(u) \\ \text{s.t} \quad & c_j(y_j, u) = A(u)y_j - q_j = 0 \quad j = 1, \dots, N \end{aligned} \quad (1)$$

[0025] For example, parameter estimation may be performed to estimate values for conductivity or susceptibility of a target, such as an oil deposit. The domain of interest may be a certain volume of land. The measured field, y , used to estimate u may be an electromagnetic field. The electromagnetic field may be generated using multiple current-injected sources located within boreholes in the land, and the electromagnetic field may be detected using receivers located on the surface of the land, for example. The electromagnetic field is affected if it intersects an oil deposit. By estimating values for the conductivity/susceptibility model from measured values for the electromagnetic field, it is accordingly possible to determine where the oil deposit is by identifying the appropriate fluctuations in the conductivity/susceptibility model. Similar methods can be employed in other fields, such as in medicine to image bones by estimating elastic parameters, to obtain images of certain media of interest.

[0026] Conventionally, Equation (1) is solved deterministically. “Deterministically” in this context refers to solving Equation (1) using algorithms that behave predictably; i.e., given a particular input, a deterministic algorithm will always produce the same output. In contrast, the embodiments described herein instead solve Equation (1) stochastically. “Stochastically” in this context refers to solving Equation (1) using algorithms that behave, at least to a certain degree, randomly. Solving Equation (1) stochastically as opposed to deterministically can result in a technical benefit of significant savings in computational resources (e.g.: savings in one or both of FLOPS and memory used). Compared to conventional deterministic methods of parameter estimation, the stochastic embodiments described herein can be about ten times faster and also typically consume significantly less memory. This results in savings in any one or more of time, hardware (processing and memory requirements), and money. For example, when performing a single instance of parameter estimation in the oil and gas industry, the embodiments described herein can result in savings of roughly \$100,000 to \$1,000,000 relative to conventional methods.

[0027] Referring now to Figure 1, there is depicted a system 100 for determining an estimated model parameter from a measured field; the system 100 can be used, for example, to estimate DC resistivity. The system 100 includes a surveying truck 102 that is stationed over a particular volume of land to be surveyed. One goal of the system 100 is to be able to locate the presence of the medium 114 located within the land. Located in boreholes within the land are sources 104 in the form of permanent current electrodes. Also positioned in the land is a receiver 106 in the form of a cone-mounted potential electrode. The cone-mounted potential electrode records voltage differences relative to a potential electrode 105, located on or near the surface of the land. Both the sources 104 and the receiver 106 are communicatively coupled to, and controlled via, the truck 102. The system 100 is accordingly configured to generate and to measure a field in the form of a static electric field. The static electric field intersects with the medium 114, which affects the field and which accordingly allows the medium 114 to be identified.

[0028] Measurements obtained via the truck 102 are sent to a processor 108 for analysis, where a log conductivity model for the volume of land can be estimated. The processor 108 is communicatively coupled to a memory 110, which contains statements and instructions to cause the processor 108 to perform a method for parameter estimation, such as the one described with

reference to Figure 2, below. A display 112 is also communicatively coupled to the processor 108 to allow the processor 108 to resolve and output an image identifying the location of the medium 114.

[0029] In contrast to attempting to solve Equation (1) deterministically, the embodiments described herein solve Equation (1) stochastically; Equation (1) is cast as a stochastic optimization problem involving estimation of the trace. To arrive at this solution, Equation (1), unconstrained, can first be re-written as:

$$\min_u J(u) = \frac{1}{2} \|PA(u)^{-1}Q - D\|_F^2 + R(u) \quad (2)$$

where $Q = (q_1, \dots, q_N)$, $D = (d_1, \dots, d_N)$, and $\|\cdot\|_F$ denotes the Frobenius norm. The difficulty is to evaluate the residual term $S(u) = PA(u)^{-1}Q - D$ and its norm. To address this difficulty, a stochastic interpretation of the trace can be used.

[0030] For a random variable/vector w with 0 mean and covariance matrix I , the following holds:

$$\mathbf{E}_w (w^\top S(u)^\top S(u) w) = \text{trace}(S(u)^\top S(u)) = \|S(u)\|_F^2. \quad (3)$$

[0031] Using Equation (3) the deterministic problem of Equation (2) can be replaced with a stochastic optimization problem, as follows:

$$\min_u \mathbf{E}_w J(u, w) = \frac{1}{2} \|(PA(u)^{-1}Q - D)w\|^2 + R(u) \quad (4)$$

[0032] In Equation (4), $\mathbf{E}_w$ is the expected value with respect to w . While Equation (4) is mathematically identical to Equation (2), it requires fewer computational resources to solve. Because Equation (4) is in the form of a “stochastic optimization problem”, a variety of algorithms referred to herein as “stochastic optimization algorithms” can be used to solve Equation (4). As long as the evaluation of the expectation in Equation (4) can be done efficiently and with sufficient accuracy, there is an efficiency gain over using deterministic algorithms.

[0033] In the stochastic formulation, each realization w_j of w involves a single solve for the partial differential equation; i.e.:

$$(PA(u)^{-1}Q - D)w_j = PA(u)^{-1}(Qw_j) - Dw_j,$$

which implies that given w_j , only a single partial differential equation is solved in order to evaluate the stochastic misfit function. This is in contrast to the N partial differential equations to be solved when evaluation Equation (2). Accordingly, relatively efficient algorithms and methods can be employed.

[0034] In particular, two end members of stochastic optimization algorithms can be considered: Stochastic Average Approximation (SAA) and Stochastic Optimization (SA).

[0035] When using SAA, one approximates the expected value by a Monte-Carlo approximation using the sum:

$$E_w J(u, w) \approx \frac{1}{K} \sum_{j=1}^K J(u, w_j).$$

$K \ll N$ different realizations are used to approximate the expected value.

[0036] When using SA, a single realization is used and the gradient $\nabla_u J(u, w_j)$ is determined. A step in the negative direction of the gradient is then taken, averaged with the previous steps and the process repeats. For each step of the SA embodiment, only a single realization is used. Applying this to solving Equation (4) implies that only a single partial differential equation and adjoint are solved at each iteration, rather than $2N$ in Equation (2) and $2K$ when using SAA. Both SA and SAA can lead to increased performance relative to conventional methods because each approximation may require fewer partial differential equation solves. Applying both SA and SAA are considered in more detail, below.

Solution Through SAA

[0037] When using SAA, the expectation can be approximated as follows:

$$\min_u \mathbf{E}_w (J_w(u, w)) \approx \hat{J}_K(u; w_1, \dots, w_K) := \frac{1}{2K} \sum_{j=1}^K \|PA(u)^{-1}Qw_j - Dw_j\|^2 + R(u). \quad (5)$$

Even with $K = 1$, the evaluation of the trace has been found to be sufficiently accurate in some contexts.

[0038] In order to choose the vectors w_j stochastic trace estimators are considered. If w is 0 mean with covariance I then $\mathbf{E}_w w^T S^T S w = \text{trace}(S^T S)$. The expected value is identical for all w 's drawn from any distribution with 0 mean and covariance I . However, as approximations are being performed, the goal is to obtain this mean with a relatively small variance, and in certain embodiments the smallest variance possible; this can help in generating a relatively small, and in certain embodiments the smallest, number of right hand sides in the partial differential equation. This can be obtained when w is a random vector drawn from the distribution of ± 1 . w is accordingly chosen from this distribution. In alternative embodiments, other choices of w are possible, although they are suboptimal in that their variance is higher.

[0039] SAA beneficially separates the questions of approximating the expectation and the optimization algorithm to be used. In order to solve the optimization problem any robust algorithm can be used. The LBFGS method is applied to the unconstrained problem in the embodiments explicitly discussed herein; however, in alternative embodiments, other methods may be applied to solve the problem.

Solution through SA

[0040] An exemplary algorithm in pseudocode that can be used for SA is given below:

- 1: Initialize the solution $u = u_0$
- 2: while not converge do
- 3: choose a realization w_k
- 4: approximately solve the optimization problem $s = \arg \min_{\delta u} J(u_k + \delta u, w_k)$

- 5: set $\hat{u}_{k+1} = u_k + \gamma s$
- 6: average $u_{k+1} = [1/(k+1)](\sum_j u_j + \hat{u}_{k+1})$
- 7: end while

[0041] Considering the foregoing, two questions arise: first, how to pick an appropriate realization w_k ; and second, how to approximately solve the optimization problem and choose the step size γ in step 5 of the foregoing algorithm. As discussed above, a beneficial choice for w is obtained by choosing w to be a random vector of $\{+/-1\}$. For an optimization algorithm, a steepest descent type method may be used, and potentially Gauss-Newton and LBFGS methods as well. As above, to facilitate comparison with SAA results, the exemplary embodiments discussed herein utilize the LBFGS method.

The Quality of the Approximation

[0042] The quality of the approximation discussed above depends on the stochastic approximation of Equation (3). For statistical problems which are set appropriately, convergence is guaranteed as the number of samples goes to infinity. Since Monte Carlo methods are used herein, convergence is typically related to the square root of the number of samples. The error is therefore associated with the fact that the problem is approximated using a finite number of sampling points and the variance of the solution is accordingly the relevant quantity to examine. To illustrate this principle, an exemplary DC resistivity problem is used. In such a DC resistivity problem, the matrix A results from discretization of the problem:

$$\nabla \cdot \exp(u) \nabla y_j = q_j \quad \nabla y_j \cdot \vec{n} = 0 \quad \int_{\Omega} y_j = 0 \quad j = 1, \dots, N$$

where the goal is to recover log conductivity u given measurements of the potential fields y_j , which result from the sources q_j . Modern acquisition systems can have hundreds and thousands of the sources q_j spread over a physical grid and in boreholes. The partial differential equation shown in Equation (1) results from performing a finite volume or finite element discretization of the differential equation.

[0043] The goal is to approximate the “misfit”:

$$\text{misfit}(u) = \|PA(u)^{-1}Q - D\|_F^2$$

using stochastic approximation. To begin, u is set to be constant, e.g. 10^{-2} , the gradient of the misfit, $s = \nabla_u \text{misfit}$, is computed, and a 1D function is defined as follows:

$$f(\alpha) := \text{misfit}(u + \alpha s) = \|PA(u + \alpha s)^{-1}Q - D\|_F^2.$$

[0044] The following approximation is then used:

$$\widehat{f}_K(\alpha) = \widehat{\text{misfit}}(u + \alpha s; w_1, \dots, w_K) = \frac{1}{K} \sum_{j=1}^K \|(PA(u + \alpha s)^{-1}Q - D)w_j\|^2.$$

[0045] Presuming 500 of the sources 104, the partial differential equation is discretized using 32^3 cells, where y and u are discretized on cell-centers. Thus, $Q : \mathbf{R}^{500} \rightarrow \mathbf{R}^{32^3}$ and to compute the misfit in the original formulation, $A(u)$ is inverted 500 times. Random realizations $[w_1, \dots, w_K]$, where $K = 1, 10, 25, 50$ are then experimented with. Evaluating the misfit in this way is roughly a factor of $500K^{-1}$ cheaper when comparing floating point operations compared to the original evaluation. Figures 3(a) – (d) are plots of the estimated misfit $\widehat{\text{misfit}}$ as a function of α . As each approximation depends on a particular choice of w that is random, five different realizations of w are used for each approximation. Although the functions are not identical, when using a single vector to evaluate the misfit, the minimum with respect to α is not far from the minimum obtained by exact evaluation of the misfit.

A Quantitative Evaluation

[0046] In the following discussion, $w_{[1,K]}$ refers to a sample of K vectors $[w_1, \dots, w_K]$. $W_{[1,K]}^l$ means the l^{th} sample of K vectors. u^* is the minimizer of the approximate objective function. Accordingly, for a given $w_{[1,K]}$:

$$u^* = \arg \min_u \hat{J}_K(u; w_{[1,K]}) = \frac{1}{K} \sum_j J(u; w_j).$$

[0047] Let $u^\dagger$ be the minimum of the exact objective function:

$$u^\dagger = \arg \min_u J(u) = \mathbf{E}_w(J_w(u, w))$$

and

$$J(u^\dagger) \leq J(u^*).$$

[0048] The quality of the approximation can be measured by the gap:

$$\text{gap} = J(u^*) - J(u^\dagger),$$

which can be estimated statistically.

[0049] To estimate $J(u^*)$, $w_{[1,K]}$ is resampled and the value of $\hat{J}(u^*, w_{[1,K]})$ is computed for different realizations of $w_{[1,K]}$. If $w_{[1,K]}$ is sampled L times, the mean and the variance can be estimated, obtaining:

$$J(u^*) \approx \frac{1}{L} \sum_{\ell} \hat{J}_K(u^*; w_{[1,K]}^{\ell}) = J_L^*.$$

An optimization problem does not need to be solved to compute this estimate.

[0050] u can then be fixed to equal u^* and the samples $w_{[1,K]}$ can be changed, and given the realizations the variance of $\hat{J}(u^*, w_{[1,K]})$ can also be estimated:

$$\sigma^* = \frac{1}{L} \sum_{\ell} (\hat{J}_K(u^*; w_{[1,K]}^{\ell}) - J_L^*)^2.$$

[0051] The combination of the mean and the variance provides a lower and an upper estimate on $J(u^*)$. In particular, the 66% confidence interval is given by:

$$J_L^* - \sigma^* \leq J(u^*) \leq J_L^* + \sigma^*. \quad (6)$$

[0052] More computational work is performed to evaluate $J(u^\dagger)$. Consider, as before, independent and identically distributed samples $w_{[1,K]}^l$, $l = 1, \dots, T$, and solve T optimization problems:

$$u_\ell^* = \arg \min_u \hat{J}_K(u; w_{[1,K]}^\ell), \quad \ell = 1, \dots, T.$$

[0053] T different solutions $[u_1^*, \dots, u_T^*]$ are obtained. The variance of the function $J(u^\dagger)$ is determined as opposed to its exact value. Using the solutions $[u_1^*, \dots, u_T^*]$ the mean and variance of $J^\dagger$ can be determined. First, let

$$J_T^\dagger = \frac{1}{T} \sum_j \hat{J}_K(u_j^*; w_{[1,K]}^j)$$

be the mean of the function value at the different solutions, and let

$$\sigma^\dagger = \frac{1}{T} \sum_j (\hat{J}_K(u_j^*; w_{[1,K]}^j) - J_T^\dagger)^2$$

be the variance of the function of the different solutions. The estimated 66% confidence interval is then given by

$$J_T^\dagger - \sigma^\dagger \leq J(u^\dagger) \leq J_T^\dagger + \sigma^\dagger. \quad (7)$$

[0054] Combining Equations (6) and (7) allows an estimate of the quality of the solution to be obtained. Unlike evaluating $J(u^*)$ where only function evaluations are performed, evaluating $J(u^\dagger)$ involves solving some optimization problems. However, in many cases very few optimization problems are solved (e.g.: between 5 and 10), and to efficiently use computational resources these optimization problems are independent and can be solved in

parallel, where a suitable system is available. Additionally, from a computational point of view solving these optimization problems can be done quickly if the solution u^* of the first optimization problem is a starting point.

Embodiment of Different Observation Operators

[0055] The foregoing embodiments assume that different experiments represented by q_i have the same observation matrix P . Although many problems share this property this is not always the case. Accordingly, the following embodiments are directed at situations in which P differs for different experiments. Practically, this may represent a situation in which the sources 104 and the receivers 106 move while measurements are taken (e.g.: if the receivers 106 are dragged by a boat during measuring). More generally, this occurs when at least one of the receivers 106 measures no data corresponding to at least one of the sources 104. In these embodiments, the misfit is first rewritten in general as

$$\text{misfit}(u) = \|C \odot (PA(u)^{-1}Q - D)\|_F^2 = \|C \odot S(u)\|_F^2 \quad (8)$$

where as before

$$S(u) = PA(u)^{-1}Q - D$$

is the residual matrix. Here $\odot$ is the point-wise Hadamard product. The matrix C is, in general, a dense matrix. To illustrate that the foregoing is indeed a generalization, consider the following cases:

- If $C = \mathbf{1}$, that is, populated by 1s, then the previous case results in which all the sources 104 share all the receivers 106.
- If C is dense and populated by different positive numbers, typically each datum has a different standard deviation and C_{ij} is the inverse of the standard deviation of each datum.
- If C is sparse then only a subset of source-receiver configuration is chosen.

- If C is diagonal, which is a relatively extreme case, it is implied that each of the sources 104 has a single one of the receivers 106. This configuration is relatively common in applied geophysical surveys in which the sources 104 and the receivers 106 move simultaneously.

[0056] As before, the misfit is represented by the following expectation:

$$\begin{aligned} \text{misfit}(u) &= \|C \odot S(u)\|_F^2 \\ &= \mathbf{E}_w (\|(C \odot S(u))w\|^2) = \mathbf{E}_w (\|(C \odot (PA(u)^{-1}Q - D))w\|^2). \end{aligned} \quad (9)$$

[0057] Evaluating Equation (9) can be difficult. Recall that for $C = \mathbf{1}$:

$$(\mathbf{1} \odot S(u))w = S(u)w = PA(u)^{-1}(Qw) - Dw.$$

[0058] However, when C does not equal $\mathbf{1}$, the same methodology cannot be followed. The difficulty is that the Hadamard product does not associate with matrix-vector product and therefore one cannot compute the product $S(u)w$ first and only then compute the Hadamard. This provides a challenge in extending the method discussed in respect of the embodiments above, in which P is constant for all different experiments. The following discusses two different approaches for solving this problem.

The Case of Low Rank C

[0059] In some applications it is possible to decompose C into low rank matrices, or at least, to approximate C by a product small rank matrices. Consider the case where

$$C = YZ^T \quad (10)$$

where Y and Z are $n_r \times k$ and $n_s \times k$ matrices respectively with $k \ll \min(n_r, n_s)$. Using the low rank representation of C the following can be proved.

[0060] First, let R be an $n_r \times n_s$ matrix with columns R_i , $i = 1, \dots, n_s$ and let Y and Z be $n_r \times k$ and $n_s \times k$ matrices respectively, with columns Y_j and Z_j , $j = 1, \dots, k$. Finally, let w be a vector of size n_s . Then:

$$(C \odot R)w = (YZ^T \odot R)w = \sum_{i=1}^{n_r} w_i R_i \odot \sum_{j=1}^k Y_j Z_{ji} = \sum_{j=1}^k \sum_{i=1}^{n_r} w_i R_i \odot Y_j Z_{ji} =$$

$$\sum_{j=1}^k \left(Y_j \odot \sum_{i=1}^{n_r} R_i Z_{ji} w_i \right) = \sum_{j=1}^k Y_j \odot R(Z_j \odot w)$$

[0061] The foregoing implies that it is possible to obtain a stochastic representation to the misfit of the form

$$\frac{1}{2} \|C \odot (P^T A(u)^{-1} Q - D)\|_F^2 = \mathbf{E}_w \left\| \sum_{j=1}^k Y_j \odot (P^T A(u)^{-1} Q - D)(Z_j \odot w) \right\|^2 =$$

$$\mathbf{E}_w \left\| \sum_{j=1}^k Y_j \odot (P^T A(u)^{-1} Q(Z_j \odot w) - D(Z_j \odot w)) \right\|^2 .$$

(11)

[0062] The above implies that it is possible to obtain a stochastic approximation similar to the case that $C = ee^T$ using k matrix solves where k is the rank of C . In cases that C is not low rank, it is possible to approximate C by a low rank matrix and use the above decomposition as a cheap alternative to the original misfit function.

The General Case

[0063] A number of different ways exist to obtain a stochastic approximation to the trace in this case. Two such approximations are discussed in the following.

[0064] One way is based on direct estimation of the product $C \odot S(u)$. It is straightforward to observe that if w is a random variable with 0 mean and identity covariance then

$$C \odot S(u) = \mathbf{E}_w (\text{diag}(S(u)w) C \text{diag}(w)) \quad (12)$$

and therefore one arrives to the following stochastic representation of the misfit:

$$\text{misfit}(u) = \|C \odot S(u)\|_F^2 = \|\mathbf{E}_w(\text{diag}(S(u)w) C \text{diag}(w))\|_F^2 \quad (13)$$

and its approximation:

$$\widehat{\text{misfit}}(u) = \left\| \frac{1}{K} \sum_{j=1}^K \text{diag}(S(u)w_j) C \text{diag}(w_j) \right\|_F^2. \quad (14)$$

[0065] A difference between the approximation of Equation (14) and those discussed in respect of other embodiments discussed above is that in Equation (14), the expected value is inside the norm, and hence the expectation (approximated by the sum) is first evaluated inside and only then is the norm evaluated. Again, only the matrix vector product $S(u)w_j$ is needed in order to evaluate the misfit. Thus the number of partial differential equations to be solved is equivalent to the number of stochastic realizations.

[0066] To obtain the derivative of the approximation to the misfit,

$$\text{vec}(\text{diag}(S(u)w_j) C \text{diag}(w_j)) = C_{w_j} S(u)w_j,$$

where the matrix C_{w_j} is defined by

$$C_{w_j} = [\text{blockdiag}((C \text{diag}(w_j))_{1 \rightarrow}, (C \text{diag}(w_j))_{2 \rightarrow}, \dots, (C \text{diag}(w_j))_{K \rightarrow})]^\top,$$

where the notation $l \rightarrow$ implies the l^{th} row of the matrix. Using this notation the approximated misfit can be rewritten as

$$\widehat{\text{misfit}}(u) = \left\| \frac{1}{K} \sum_{j=1}^K C_{w_j} S(u)w_j \right\|^2 = \left\| \frac{1}{K} \sum_{j=1}^K C_{w_j} (PA(u)^{-1}Qw_j - Dw_j) \right\|^2.$$

[0067] This representation allows for simple differentiation of the approximated misfit

$$\nabla_u \widehat{\text{misfit}}(u) = - \left(\frac{1}{K} \sum_{j=1}^K G_j^\top A(u)^{-\top} P^\top C_{w_j}^\top \right) \left(\frac{1}{K} \sum_{j=1}^K C_{w_j} S(u)w_j \right),$$

where the matrix $G_j = G(Qw_j, u)$ is defined by $G_j(y_j, u) = \nabla u[A(u)y_j]$. This formulation of the problem requires K solves of the adjoint problem.

[0068] Other stochastic approximations to the trace can be obtained. Starting from Equation (9), the following identity can be applied:

$$(C \odot S(u)) w = \text{diag}(C \text{diag}(w) S(u)^\top)$$

and therefore

$$\text{misfit}(u) = \|C \odot S(u)\|_F^2 = \mathbf{E}_w (\|\text{diag}(C \text{diag}(w) S(u)^\top)\|^2). \quad (15)$$

[0069] Identity (15) does not solve the problem of computing either $A(u)^{-1}Q$ or $A(u)^{-\top}P$. The point here is that evaluating the diagonal of the matrix product

$$C \text{diag}(w) S(u)^\top = C \text{diag}(w) (PA(u)^{-1}Q - D)^\top$$

is needed without computing the matrix itself. Thus, another stochastic process is used. If x is a random variable for zero mean and covariance $\text{Cov}(x) = I$, then:

$$\text{diag}(C \text{diag}(w) S(u)^\top) = \mathbf{E}_x (x \odot (C \text{diag}(w) S(u)^\top) x). \quad (16)$$

[0070] Thus the misfit can be replaced with a double stochastic process

$$\text{misfit}(u) = \|C \odot S(u)\|_F^2 = \mathbf{E}_w (\|\mathbf{E}_x (x \odot (C \text{diag}(w) S(u)^\top) x)\|^2). \quad (17)$$

[0071] Again, using Monte-Carlo to obtain an approximation results in:

$$\widehat{\text{misfit}}(u) = \frac{1}{K_1 K_2} \sum_{i=1}^{K_1} \left\| \sum_{j=1}^{K_2} (x_j \odot (C \text{diag}(w_i) S(u)^\top x_j)) \right\|^2, \quad (18)$$

which requires the computation of $S(u)^T x$ that implies the solution of the adjoint problem. Experimentally, it has been found that in general, the approximation (18) requires more realizations than the approximation (14) in order to achieve similar accuracy of the misfit.

[0072] Using the stochastic misfit approximation $\widehat{\text{misfit}}(u)$ given by either (14) or (18) the SAA method can be used to solve the stochastic optimization problem and approximately recover the model u . Notably, since the expected value is inside the norm using SA is not justified.

Numerical Experiments

[0073] The following discusses numerical experiments performed to test the foregoing approximations and to experiment with different methods and parameters. Both the exemplary DC resistivity problem, discussed above, is used, as well as a seismic tomography/full-wave inversion problem. In the seismic tomography problem, consider the case where the matrix A results from discretization of the problem

$$(\Delta + k^2(u)) y_j = q_j, \quad \text{in } \Omega \subset \mathbb{R}^2, \quad y_j = 0 \text{ on } \Gamma = \partial\Omega, \quad j = 1, \dots, N,$$

where $k^2(u) = 2\pi f u$ is the wave number with f the temporal frequency, and $u = c^{-2}$ with c the acoustic wave-speed. In this case the goal is to recover the velocity $c = u^{-1}$ given a measurements of the wave-field fields y_j for each source experiment. Modern acquisition systems can have hundreds and thousands of sources spread over a physical grid and in boreholes.

Experiments with DC Resistivity Problem

[0074] For the DC resistivity problem, the code for the forward problem and Jacobians is based on a software package described at A. Pidlisecky, E. Haber, and R. Knight, Resinvm3d: A matlab 3-d resistivity inversion package, *Geophysics*, 72:212-224, 2006.

[0075] In particular, the following compares conventional deterministic methods with the stochastic methods of the embodiments discussed herein. Since the deterministic methods utilize

large amounts of computer memory, the LBFGS algorithm is used for all methods such that the comparison between deterministic and stochastic methods is on equal footing.

[0076] The following utilizes a mesh of 32^3 cells. For the embodiment in which all the sources 104 share the same receivers 106, the experiments are performed assuming 250 each of the sources 104 and the receivers 106. As depicted in Figure 4, which is a schematic of the system used to perform the DC resistivity experiment, the sources 104 are located in boreholes while the receivers 106 are located in a grid-like arrangement on the surface. Also as depicted in Figure 4, the media 114 are an ellipsoid conductor with a conductivity of 10 S/m and a resistor with a conductivity of 0.1 S/m. For simplicity a smooth recovery is used, and

$$R(u) = \frac{\rho}{2} u^\top \nabla_h^\top \nabla_h u,$$

where ∇_h is a discretization of the gradient operator and ρ is chosen such that a fit to the data is obtained based on the discrepancy principle.

Recovery When All the Sources 104 Share the Same Receivers 106

[0077] In a first experiment in which it is assumed that all the sources 104 share all the same receivers 106, the stochastic problem of Equation (4) results. To avoid the “inverse crime”, data for the true model is generated using a fine grid of 64^3 . Three different methods are then used to solve the inverse problem; i.e., to solve for the model u based on the measured field. In the first method, the deterministic LBFGS method is used to yield what is conventionally considered an “exact” solution of the stochastic optimization problem. In a second experiment and according to one of the embodiments, five vectors w are selected and the SAA embodiment is used to stochastically obtain a solution. In a third embodiment, the SA embodiment is used with a single realization at each step, while replacing realizations as the method progresses. Figures 5(a) – (d) show the different results: Figure 5(a) shows the true solution; Figure 5(b) shows the deterministic solution; Figure 5(c) shows the solution using SA; and Figure 5(d) shows the solution using SAA. The results obtained using the stochastic optimization methods according to the depicted embodiments closely resemble the deterministic optimization: the relative difference in the recovered models by the different techniques is below 3%. In this

particular example, it is possible to compute the true value of the misfit and its stochastic approximation. The relative difference between the misfit obtained by the deterministic optimization method and the misfit obtained by the SAA embodiment was 0.11%, and the difference between the misfit obtained by the SA embodiment and the misfit obtained by the deterministic optimization was 0.13%. Thus, the SA and SAA embodiments are able to closely approximate the true solution, while being much more computationally efficient in terms of computational resources used, as illustrated below in Table 1 by the lower number of partial differential equations solved while performing the SA and SAA embodiments:

Table 1: Number of Iterations and Partial Differential Equation Solves Used by Deterministic, SAA, and SA Embodiments

Method	Iterations	Partial Differential Equation Solves	Difference Between Solution and Deterministic Solution
Deterministic	51	25,500	0%
SAA	43	430	2.1%
SA	325	325	1.2%

[0078] Table 1 shows that a factor of about 60 in terms of partial differential equation solves is observed between the stochastic and deterministic methods. While the estimates obtained using the different methods are roughly similar, the savings in resources provided by the SA and SAA embodiments allows larger problems to be tackled using a fraction of the computational effort used when solving deterministically.

[0079] For the foregoing:

- For the deterministic and SAA optimization, computations are stopped when the norm of the gradient is 10^{-3} of its initial value.

- For the SA embodiment, LBFGS is used, and stopped when the change in the solution is smaller than 10^{-3} .
- For the SAA embodiment, five vectors are used to obtain the solution. The estimates discussed above are used as stopping criteria. The number of w 's is determined using experimentation. When evaluating the gap $= J(u^*) - J(u^\dagger)$ the problem is first solved for five different samples starting from the solution obtained by the first sample. For all problems the solutions for different samples converge in a single iteration, which indicates that iterations began close to the solution and the variability between solutions is small.

The question of how many realizations of w 's to use is determined through experimentation. From the starting point, the variability of the misfit is experimented with. For a small set of w 's (for example using a single vector), a reasonable approximation to the misfit (and a reasonable recovery of u) is obtained, but repeating the experiment a non-negligible variance of about 7% between the solutions is found from different realizations. When using five realizations the results change relatively little (less than 3%) between different realizations. Determining the number of realizations *a priori* can be beneficial.

Recovery When All the Sources 104 Do Not Share the Same Receivers 106

[0080] In a second set of experiments, the stochastic approximations given in Equations (14) and (18) are used to solve the problem where the matrix C is not equal to $\mathbf{1}$. Three different cases are tested:

- First, the matrix C is dense and contains the standard deviation of each datum. 1% noise is assumed, and thus each entry in C is roughly 1% of the value of the corresponding datum.
- Second, the matrix C is sparse, made of 0's and 1's, and contains only 10% nonzero entries. The entries are randomly chosen.

- Third, there are an equal number of the sources 104 and the receivers 106 and the matrix C is diagonal, $C = I$; i.e., each of the sources 104 has a single one of the receivers 106. This case is relatively common in certain geophysical applications in which the sources 104 and receivers 106 simultaneously move.

[0081] Three different optimization problems are solved for each of the foregoing:

- one, a deterministic LBFGS;
- two, SAA with the stochastic approximation in Equation (14) (referred to as SAA1 in Table 2, below); and
- three, SAA with the stochastic approximation in Equation (18) (referred to as SAA2 in Table 2, below).

[0082] In this second set of experiments, 500 of the sources 104 and 500 of the receivers 106 are used in the same configuration as shown in Figure 4. Again the number of iterations and partial differential equation solves are recorded for each of the cases, as well as the difference between the models u recovered by stochastic optimizations relative to the models recovered for the deterministic one. The results for the different cases are summarized in Table 2, below.

Table 2: Number of Iterations and Partial Differential Equation Solves Used by the Dense, Sparse, and Diagonal Cases

C	Method	Iterations	Partially Differential Equation Solves	Difference Between Solution and Deterministic Solution
Dense	Deterministic	25	26,500	0%
	SAA1	33	321	1.5%
	SAA2	38	654	5.1%
Sparse	Deterministic	31	33,500	0%

	SAA1	32	325	1.7%
	SAA2	36	745	5.3%
Diagonal	Deterministic	27	28,500	0%
	SAA1	32	331	0.9%
	SAA2	39	789	6.2%

[0083] In respect of the experiments generating the data for Table 2:

- For SAA1, five vectors were used and experimentally it was found that this was sufficient to obtain results with relatively little variability.
- For SAA2, two vectors were used for w_1 and five vectors for w_2 , and the computations are almost double compared to SAA1. As can be seen in Table 2, SAA1 results in a smaller variance compared to SAA2.
- All the optimization algorithms were terminated when the gradient was reduced to 10^{-3} of its original value.

[0084] As shown in Table 2, the model u recovered by statistical embodiments is similar to the models recovered by the deterministic models. Also as above, a technical benefit is that the depicted embodiments are performed in a fraction of the computational cost (e.g.: measured in FLOPS) of the deterministic methods.

Experiments with Seismic Tomography

[0085] In this experiment, the 2D seismic borehole tomography example presented above is employed. Two boreholes are assumed to be 1 meter apart and it is assumed that the left borehole contains 500 of the sources 104 while the right borehole contains 500 of the receivers 106. A frequency of 2 Hz is used that, for the material properties assumed in this experiment, implies roughly 3 wavelengths between boreholes. A sketch of the seismic tomography

experiment and the true model is plotted in Figure 6, in which the estimated model parameter is seismic velocity. The sources 104 are represented by the dark bar on the right of Figure 6, while the receivers 106 are represented by the light bar on the left of Figure 6.

[0086] In this experiment, SAA is used; however, unlike in the DC resistivity experiment where the number of vectors is set *a priori*, continuation is used to determine the number of samples needed. A single sample is used to start, the optimization problem is solved, and then two new samples are used starting from the previous solution. This is continued until the recovered model does not change much (e.g.: less than 1% change). Figures 7(a) – (f) plot the models obtained following this process using one to six samples.

[0087] As can be seen, using a single sample was not particularly beneficial; however, when using five and then six the result was relatively unchanged.

[0088] In the foregoing embodiments, although LBFGS is used, other algorithms can be used that converge relatively better and faster. For example, the Gauss-Newton method and methods based on equality constrained optimization can potentially be used and should have better numerical properties. Using better optimization techniques may further expand the gap between stochastic and deterministic optimization techniques.

[0089] Given the foregoing, and referring back to the system 100 of Figure 1, the following is a discussion of how the processor 108 handles the measurements it obtains using the receivers 106. The processor 108 performs a method 200, depicted in Figure 2, for determining an estimated model parameter from a measured field, according to another embodiment. The method begins at block 202, from which the processor 108 transitions to block 204 where it casts the inverse problem (i.e.: a full-waveform inversion problem) as a stochastic optimization problem so as to be able to apply a stochastic optimization technique, such as SA and SAA, to it as opposed to conventional deterministic algorithms. As discussed above, and as referenced in block 204, employing stochastic optimization involves minimizing an objective function that is present when the inverse problem is cast as a stochastic optimization problem. The processor 108 then proceeds to block 206, where the objective function is iteratively optimized using a stochastic optimization technique such as SA or SAA. Eventually, the objective function will be reduced to a desired value; the desired value may be user controlled, depending on the level of

computational resources available and the accuracy of the estimated model parameter desired. When the processor 108 has determined that the objective function has been reduced to the desired value, the processor 108 proceeds to block 210 where the model parameter that corresponds to the objective function reduced to the desired value is adopted, or identified as, the estimated model parameter. The estimated model parameters can optionally be used to generate an image of one of the media 114 located in the domain of interest. Once the estimated model parameter is obtained, the processor 108 proceeds to block 212, where the method 200 ends.

[0090] Any of the foregoing methods may be automated using any suitable type of processor, which may include a programmable logic controller, microprocessor, microcontroller, application specific integrated circuit, field programmable gate array, multi-core processor, an assembly of processors configured to execute code in parallel, or the like. Any of the foregoing methods may also be encoded on to a memory or other computer readable medium, which may be communicatively coupled to the processor. The memory may, for example, be any suitable type of semiconductor or disc based memory, such as RAM (whether volatile or non-volatile), ROM, hard disk drives, CD-ROMs, and DVD-ROMs.

[0091] While particular embodiments have been described in the foregoing, it is to be understood that other embodiments are possible and are intended to be included herein. It will be clear to any person skilled in the art that modifications of and adjustments to the foregoing embodiments, not shown, are possible.

CLAIMS

1. A method for determining an estimated model parameter from a measured field, the method comprising:
 - (a) casting an inverse problem relating the measured field to the model parameter as a stochastic optimization problem, wherein the stochastic optimization problem comprises minimizing an objective function comprising a misfit expressed as an expected value of a function that depends on a random variable;
 - (b) iteratively optimizing the objective function using a stochastic optimization technique;
 - (c) determining when the objective function is reduced to a desired value; and
 - (d) when the objective function is reduced to the desired value, identifying a model parameter that corresponds to the objective function at the desired value as the estimated model parameter.
2. A method as claimed in claim 1 wherein the measured field intersects a medium, and further comprising reconstructing an image of the medium from the estimated model parameter.
3. A method as claimed in any one of claims 1 and 2 further comprising generating the measured field using a source, and measuring the measured field using a receiver.
4. A method as claimed in claim 3 wherein the source comprises multiple sources and the receiver comprises multiple receivers, and wherein at least one of the receivers does not measure data from at least one of the sources due to relative motion between the at least one of the receivers and the at least one of the sources.
5. A method as claimed in any one of claims 1 to 4 wherein the stochastic optimization technique comprises sample average approximation.

6. A method as claimed in any one of claims 1 to 4 wherein the stochastic optimization technique comprises stochastic approximation.
7. A method as claimed in any one of claims 1 to 6 wherein the measured field is generated in response to a current signal, the field is a static electric field, and the model parameter is conductivity.
8. A method as claimed in any one of claims 1 to 6 wherein the measured field is generated in response to either a current signal or a magnetic field, the field is a magnetic field, and the model parameter is conductivity.
9. A method as claimed in any one of claims 1 to 6 wherein the measured field is generated in response to an acoustic signal, the field is either a pressure field or a displacement field, and the model parameter is seismic velocity.
10. A system for determining an estimated model parameter from a measured field, the system comprising:
 - (a) a processor; and
 - (b) a memory communicatively coupled to the processor and having encoded thereon statements and instructions to cause the processor to execute a method comprising:
 - (i) casting an inverse problem relating the measured field to the model parameter as a stochastic optimization problem, wherein the stochastic optimization problem comprises minimizing an objective function comprising a misfit expressed as an expected value of a function that depends on a random variable;
 - (ii) iteratively optimizing the objective function using a stochastic optimization technique;
 - (iii) determining when the objective function is reduced to a desired value; and

- (iv) when the objective function is reduced to the desired value, identifying a model parameter that corresponds to the objective function at the desired value as the estimated model parameter.
- 11. A system as claimed in claim 10 further comprising a display communicatively coupled to the processor, and wherein the method further comprises reconstructing an image of a medium that the measured field intersects from the estimated model parameter.
- 12. A system as claimed in any one of claims 10 and 11 further comprising:
 - (a) a source, communicatively coupled to the processor and located at a source position, configured to emit the measured field when stimulated using a signal; and
 - (b) a receiver, communicatively coupled to the processor and located at a receiver position separated from the source position, configured to measure the measured field.
- 13. A system as claimed in claim 12 wherein the source comprises multiple sources and the receiver comprises multiple receivers, and wherein at least one of the receivers does not measure data from at least one of the sources due to relative motion between the at least one of the receivers and the at least one of the sources.
- 14. A system as claimed in any one of claims 10 to 13 wherein the stochastic optimization technique comprises sample average approximation.
- 15. A system as claimed in any one of claims 10 to 13 wherein the stochastic optimization technique comprises stochastic approximation.
- 16. A system as claimed in any one of claims 10 to 15 wherein the measured field wherein the measured field is generated in response to a current signal, the field is a static electric field, and the model parameter is conductivity.

17. A system as claimed in any one of claims 10 to 15 wherein the measured field is generated in response to either a current signal or a magnetic field, the field is a magnetic field, and the model parameter is conductivity.
18. A system as claimed in any one of claims 10 to 15 wherein the measured field is generated in response to an acoustic signal, the field is either a pressure field or a displacement field, and the model parameter is seismic velocity.
19. A computer readable medium having encoded thereon statements and instructions to cause a processor to execute a method as claimed in any one of claims 1 to 9.

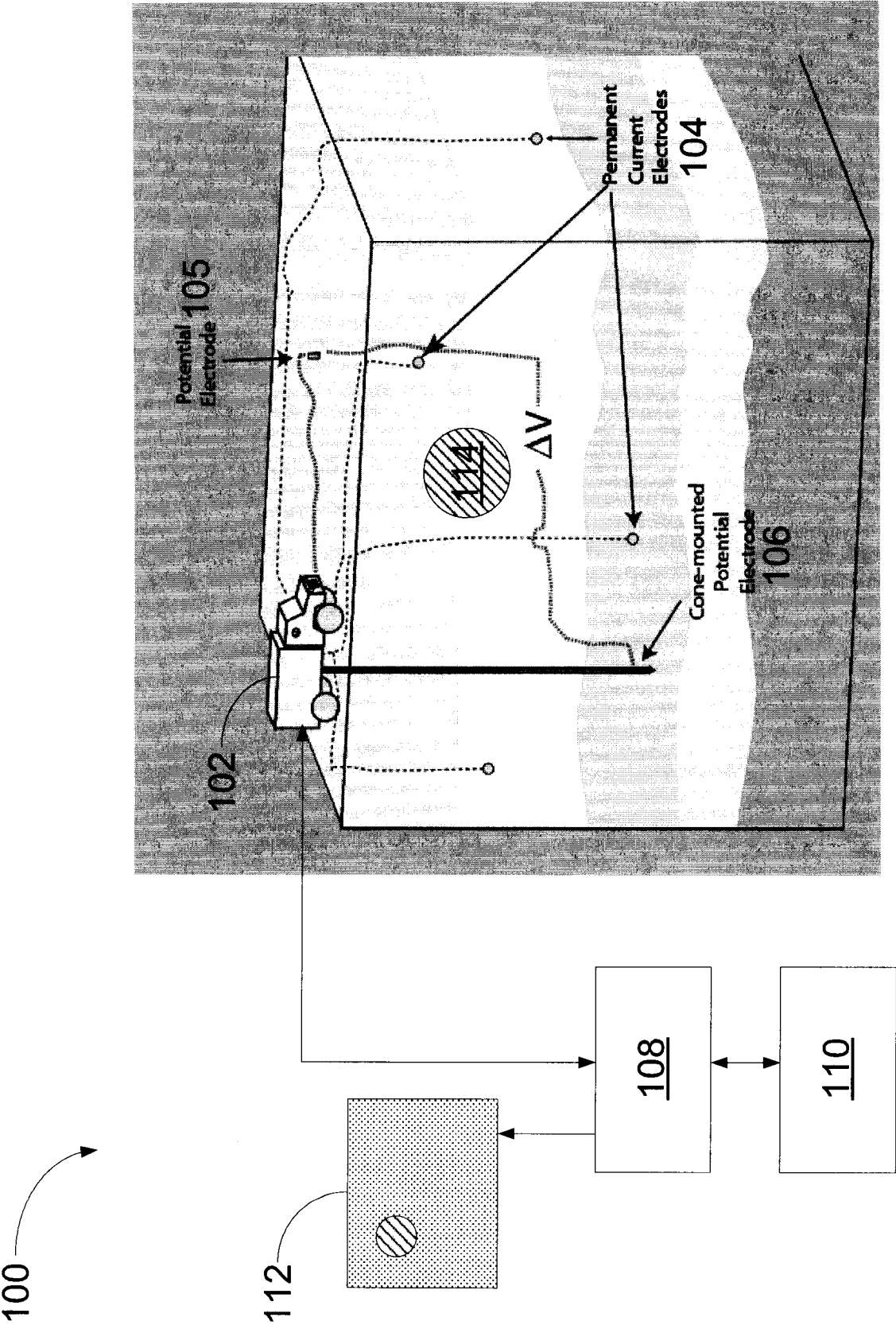
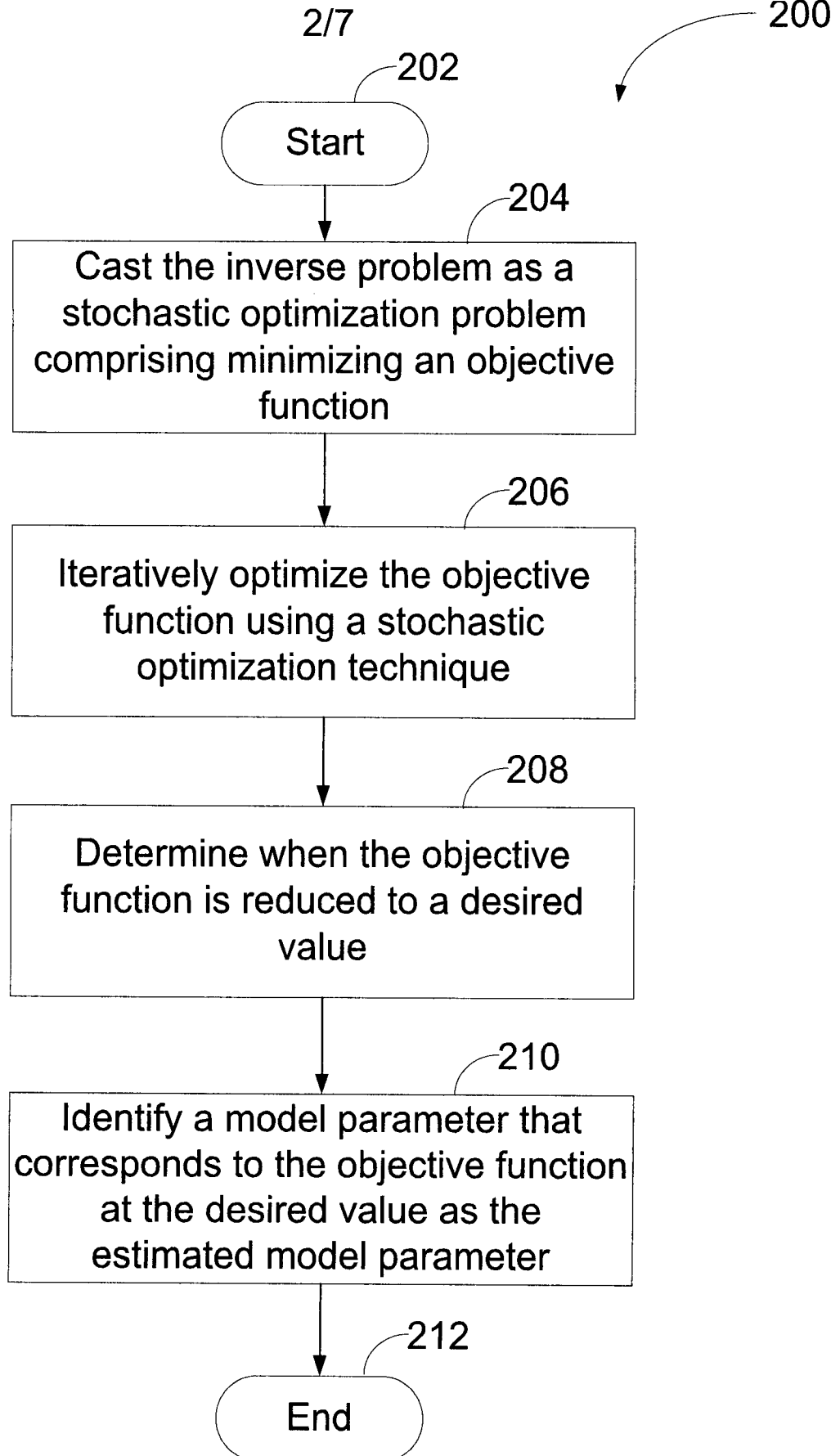


FIG. 1

**FIG. 2**

1 vector

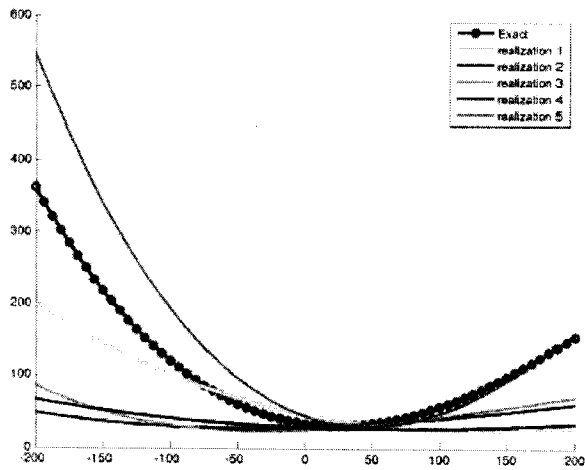


FIG. 3(a)

10 vectors

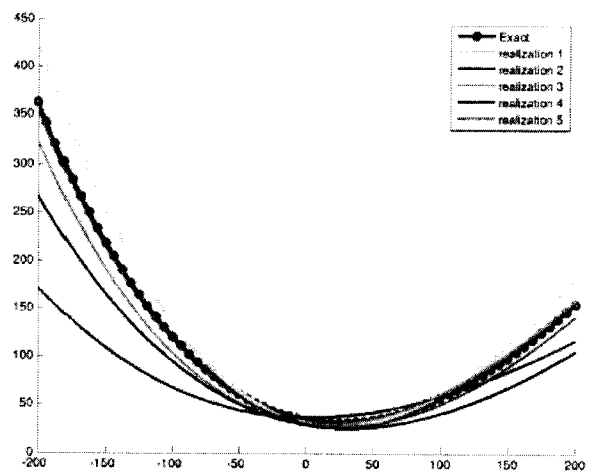


FIG. 3(b)

25 vectors

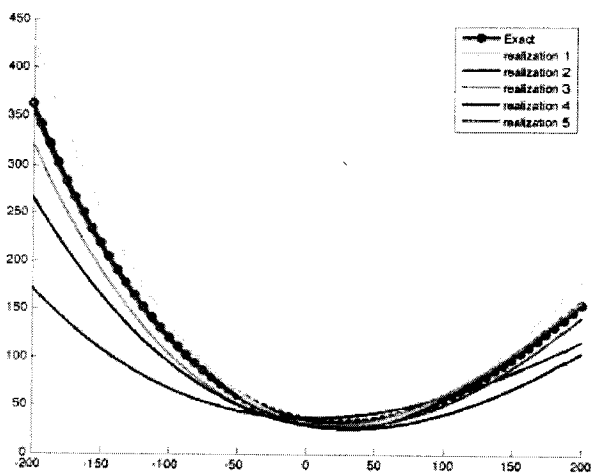


FIG. 3(c)

50 vectors

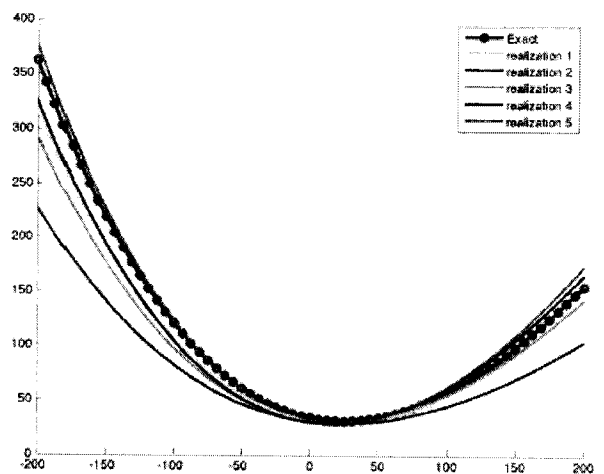


FIG. 3(d)

4/7

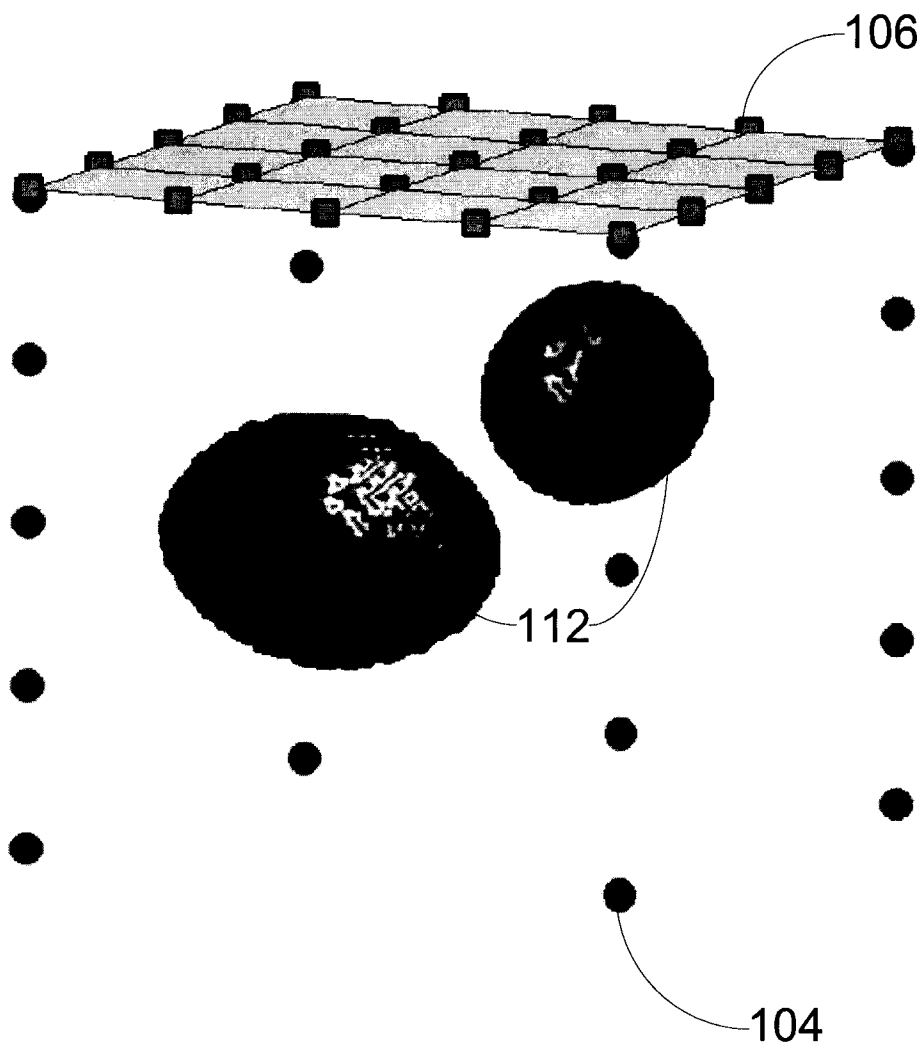
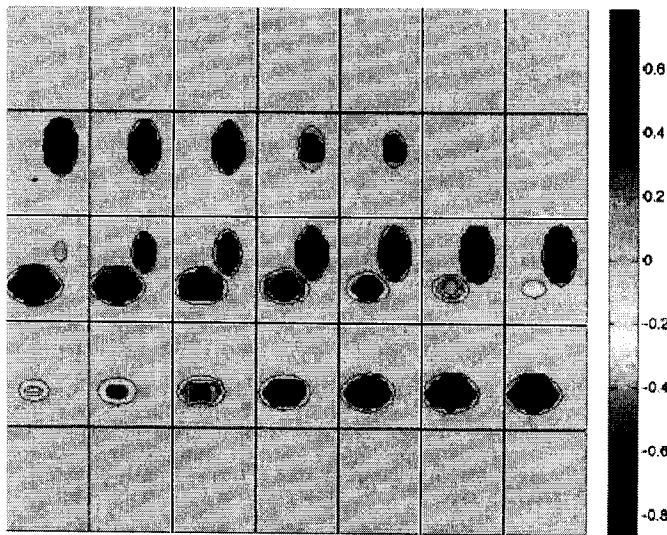
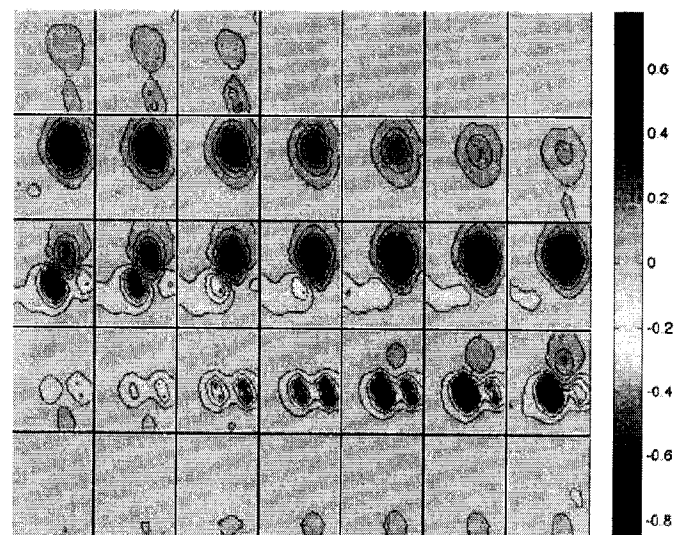


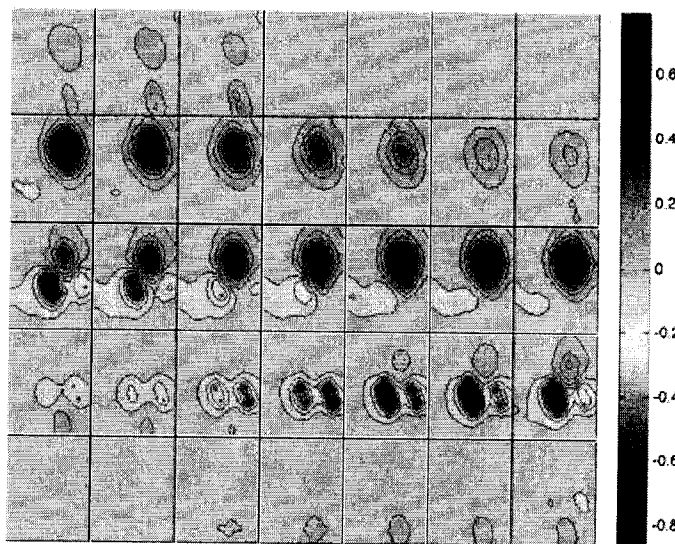
FIG. 4



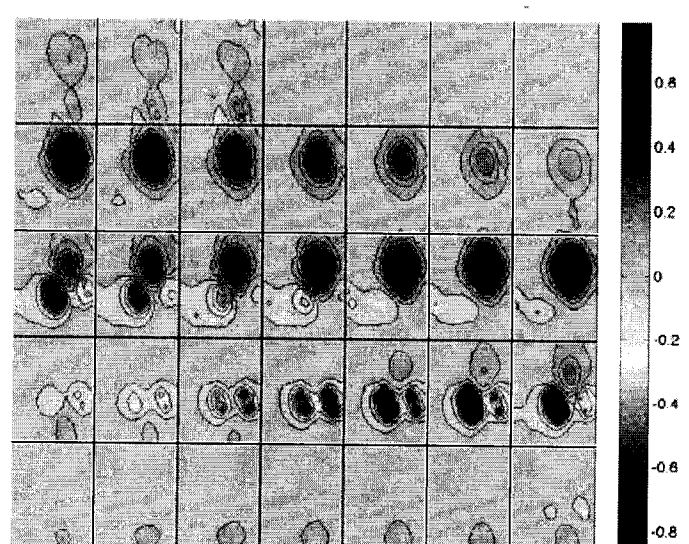
True solution

FIG. 5(a)

Deterministic recovery

FIG. 5(b)

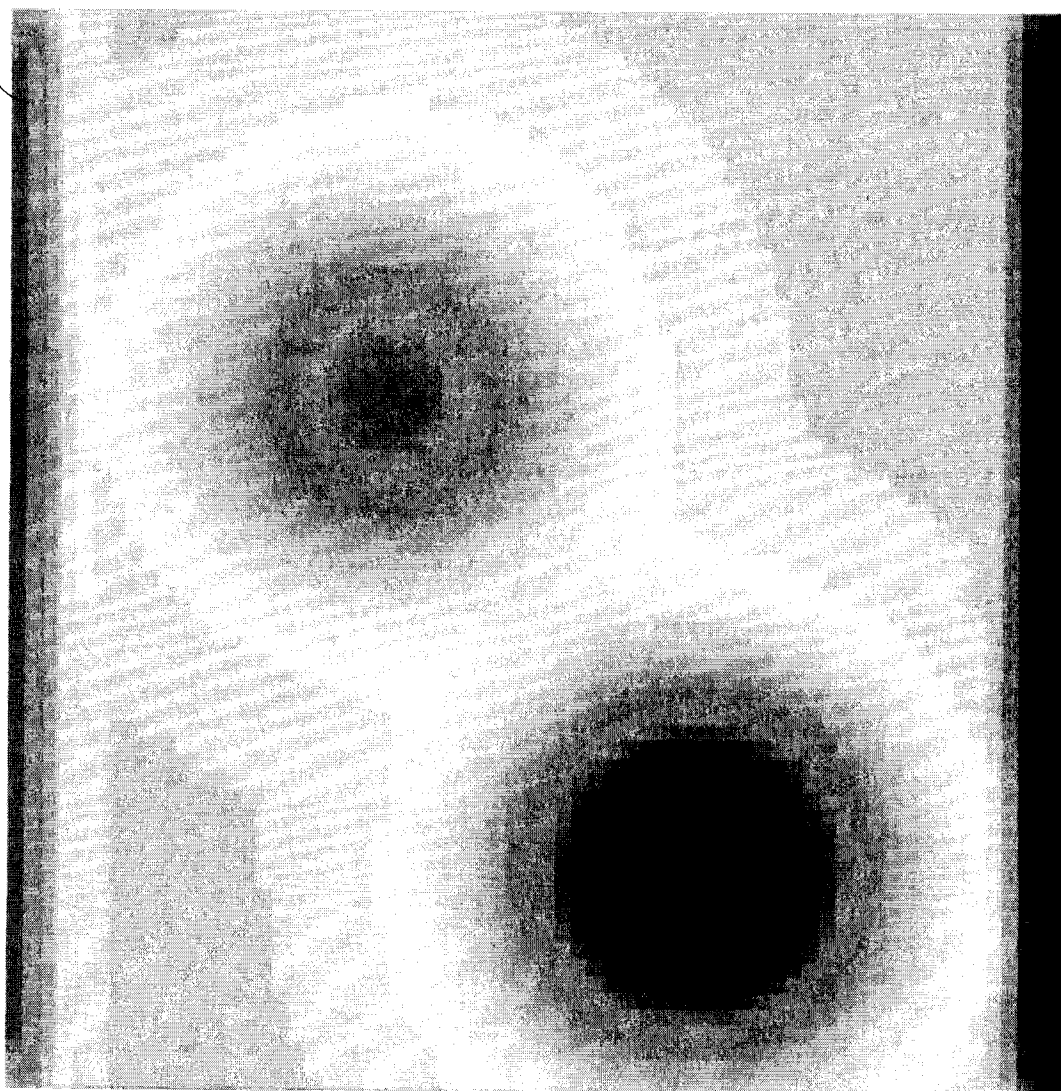
SA

FIG. 5(c)

SAA

FIG. 5(d)

106



104

FIG. 6

1 vec

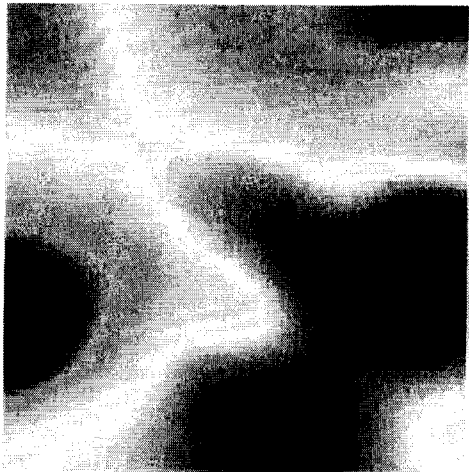


FIG. 7(a)

2 vecs

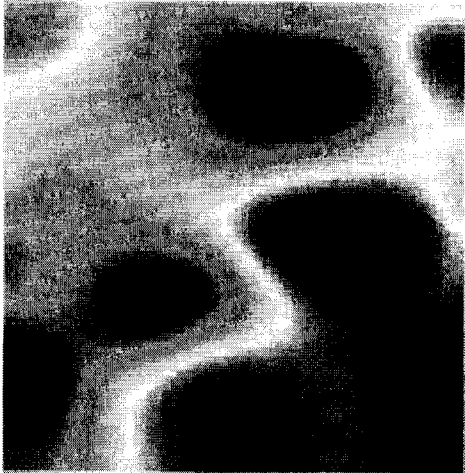


FIG. 7(b)

3 vecs



FIG. 7(c)

4 vecs



FIG. 7(d)

5 vecs

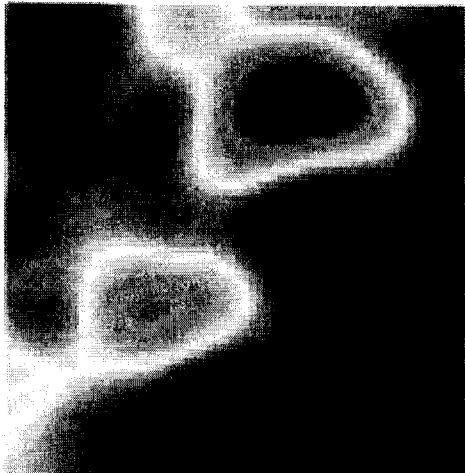


FIG. 7(e)

6 vecs

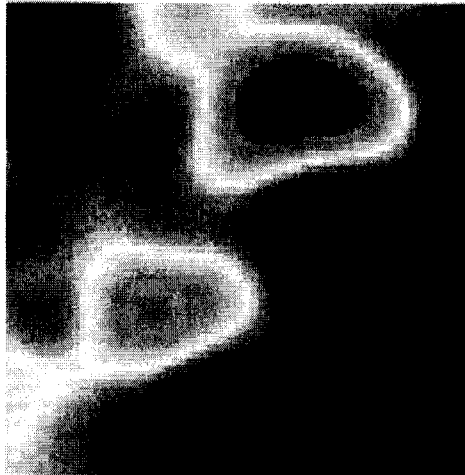


FIG. 7(f)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2011/000707

A. CLASSIFICATION OF SUBJECT MATTER IPC: G06F 17/10 (2006.01) , G01R 27/00 (2006.01) , G01V 1/30 (2006.01) , G01V 3/38 (2006.01) , G06F 17/11 (2006.01) According to International Patent Classification (IPC) or to both national classification and IPC														
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC: G06F 17/10 (2006.01) , G01R 27/00 (2006.01) , G01V 1/30 (2006.01) , G01V 3/38 (2006.01) , G06F 17/11 (2006.01) in combination with keywords Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used) Canadian Patent Database, EPODOC, TXTE, Google Scholar Keywords: model, parameter, estimation, stochastic optimization														
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>HABER, E. et al.: "Inversion of 3d Electromagnetic Data in Frequency and Time Domain Using an Inexact All-at-once Approach", Geophysics, vol. 69, no. 5, September-October 2004, pp. 1216-1228. See the entire document.</td> <td></td> </tr> <tr> <td>A</td> <td>NEMIROVSKI, A. et al.: "Robust Stochastic Approximation Approach to Stochastic Programming", SIAM Journal on Optimization, vol. 19, no. 4, pp. 1574-1609 See the entire document.</td> <td></td> </tr> <tr> <td>A</td> <td>US20040138862 A1 (HU, Lin-Ying; et al.) 15 July 2004 (15-07-2004) See abstract, paragraphs 1-8, 15-23, 30-37.</td> <td></td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	HABER, E. et al.: "Inversion of 3d Electromagnetic Data in Frequency and Time Domain Using an Inexact All-at-once Approach", Geophysics, vol. 69, no. 5, September-October 2004, pp. 1216-1228. See the entire document.		A	NEMIROVSKI, A. et al.: "Robust Stochastic Approximation Approach to Stochastic Programming", SIAM Journal on Optimization, vol. 19, no. 4, pp. 1574-1609 See the entire document.		A	US20040138862 A1 (HU, Lin-Ying; et al.) 15 July 2004 (15-07-2004) See abstract, paragraphs 1-8, 15-23, 30-37.	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
A	HABER, E. et al.: "Inversion of 3d Electromagnetic Data in Frequency and Time Domain Using an Inexact All-at-once Approach", Geophysics, vol. 69, no. 5, September-October 2004, pp. 1216-1228. See the entire document.													
A	NEMIROVSKI, A. et al.: "Robust Stochastic Approximation Approach to Stochastic Programming", SIAM Journal on Optimization, vol. 19, no. 4, pp. 1574-1609 See the entire document.													
A	US20040138862 A1 (HU, Lin-Ying; et al.) 15 July 2004 (15-07-2004) See abstract, paragraphs 1-8, 15-23, 30-37.													
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.														
<table border="0"> <tr> <td>* Special categories of cited documents :</td> <td>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td> </tr> <tr> <td>"A" document defining the general state of the art which is not considered to be of particular relevance</td> <td>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td> </tr> <tr> <td>"E" earlier application or patent but published on or after the international filing date</td> <td>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td> </tr> <tr> <td>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td> <td>"&" document member of the same patent family</td> </tr> <tr> <td>"O" document referring to an oral disclosure, use, exhibition or other means</td> <td></td> </tr> <tr> <td>"P" document published prior to the international filing date but later than the priority date claimed</td> <td></td> </tr> </table>			* Special categories of cited documents :	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family	"O" document referring to an oral disclosure, use, exhibition or other means		"P" document published prior to the international filing date but later than the priority date claimed	
* Special categories of cited documents :	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention													
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone													
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art													
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family													
"O" document referring to an oral disclosure, use, exhibition or other means														
"P" document published prior to the international filing date but later than the priority date claimed														
Date of the actual completion of the international search 31 August 2011 (31-08-2011)		Date of mailing of the international search report 6 September 2011 (06-09-2011)												
Name and mailing address of the ISA/CA Canadian Intellectual Property Office Place du Portage I, C114 - 1st Floor, Box PCT 50 Victoria Street Gatineau, Quebec K1A 0C9 Facsimile No.: 001-819-953-2476		Authorized officer Corneliu Remes (819) 934-2675												

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CA2011/000707

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
US2004138862A1	15 July 2004 (15-07-2004)	CA2445858A1 FR2846767A1 FR2846767B1 GB0325132D0 GB2394813A GB2394813B NL1024620A1 NL1024620C2 NO20034830D0 NO20034830A US7200533B2	30 April 2004 (30-04-2004) 07 May 2004 (07-05-2004) 24 December 2004 (24-12-2004) 03 December 2003 (03-12-2003) 05 May 2004 (05-05-2004) 12 July 2006 (12-07-2006) 06 May 2004 (06-05-2004) 14 September 2006 (14-09-2006) 29 October 2003 (29-10-2003) 03 May 2004 (03-05-2004) 03 April 2007 (03-04-2007)