Determination of the Mechanical Properties of a Solid Elastic Medium from a Seismic Wave Propagation Using Two Statistical Estimators

SAMIH ZEIN
ÉDOUARD CANOT
JOCELYNE ERHEL
INRIA, Campus de Beaulieu 35042, Rennes, France

NABIL NASSIF
Department of Mathematics, American University of Beirut, Riad El-Solh, Beirut, Lebanon

Abstract: In this paper, we study an inverse problem consisting in the determination of the mechanical properties of a layered solid elastic medium in contact with a fluid medium by measuring the variation of the pressure in the fluid while propagating a seismic/acoustic wave. The estimation of mechanical parameters of the solid is obtained from the simulation of a seismic wave propagation model governed by a system of partial differential equations.

Two stochastic methods, Markov chain Monte Carlo with an accelerated version and simultaneous perturbation stochastic approximation, are implemented and compared with respect to cost and accuracy.

1. INTRODUCTION

Marine geosciences study the genesis and the dynamics of processes taking place at the ocean–subsoil interface and the neighboring solid sublayers. Such studies provide a deeper knowledge of the impact of marine processes resulting from petroleum industries on environment and natural resources. To improve our knowledge of these processes, it is necessary to have equipment capable of recognition of the subsoil marine geological properties. This is done by seismic campaigns that send punctual waves then measure the reflection of these waves on each geological layer.
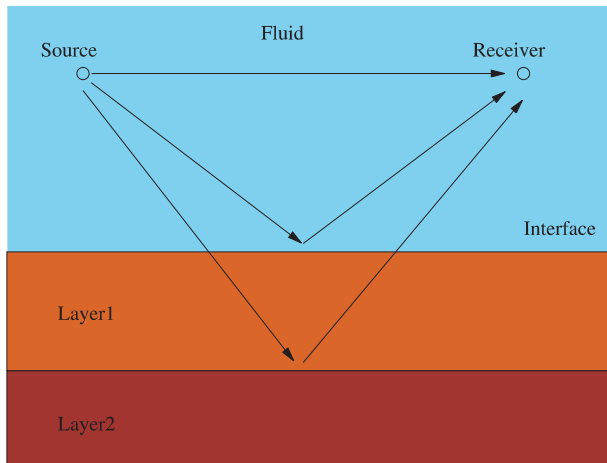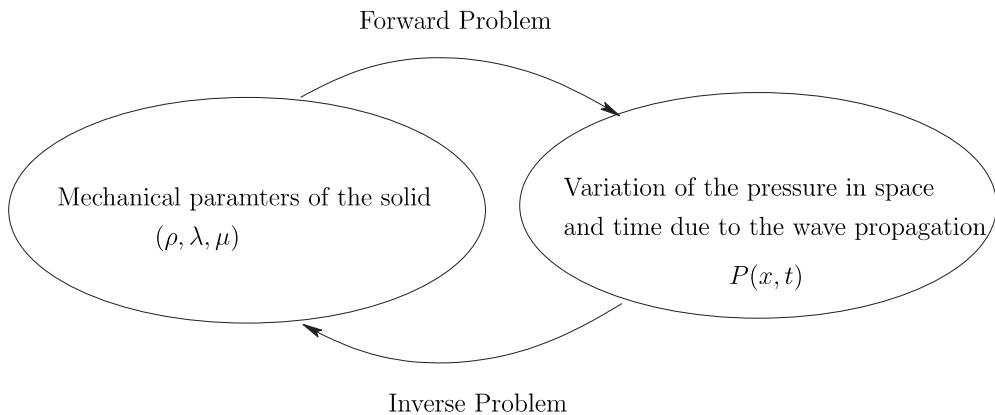
Figure 1. Seismic wave path.

Figure 2. Definition of the forward and the inverse problems.

In this paper, we consider a solid elastic medium composed of a stack of thin layers and in contact with a fluid. A numerical simulation is done by placing a seismic source in the fluid and computing the propagation of a seismic wave in both media. The seismic wave is created wihle propagating a local variation of the pressure in the fluid and a local stress in the solid. The wave is supposed to penetrate in the solid and return to the fluid after some reflections on the solid interfaces (see Figure 1). A simulation of a seismic recording at a defined receiver in the fluid shows the variation of the pressure in time at the position of the receiver. This variation depends on the mechanical properties of the solid. We are interested in finding these mechanical properties, namely the density and the Lamé's coefficients, from the pressure variation (in space and time) in the fluid due to the propagation of a seismic/acoustic wave. To solve this task, two problems have to be defined (see Figure 2):

1. the forward problem consists in finding the pressure field in the fluid knowing the mechanical properties of the solid by solving numerically a system of partial derivative equations which model the wave propagation;
2. the inverse problem is going backward to find the mechanical properties of the solid given a set of pressure measures in the fluid.

Here, the solid is a stack of one, two or three thin layers which have their thickness less than the wave length. This makes the signals coming from different reflections overlap and the inverse problem harder. This parameter estimation problem has been studied with many approaches. The essential differences between these approaches are the modelization of the wave propagation and the way to find the solution of the wave equation. Numerical and analytical methods to solve this equation have been developed according to the application domain. For example:

- In geophysics where a seismic traveltime inversion is of interest to find the wave velocity in each geological layer, the wave equation is approximated with the eikonal equation and is solved numerically using the ray tracing technique (see [1]). This approach is able to find one parameter (the wave velocity is function of the mechanical parameters) out of the three parameters so each layer is not fully determined.
- In nondestructive control, inversion techniques based on analytical solutions were studied in the case of a smoothly inhomogeneous solid. The analytical solution of the wave equation is expressed in terms of Green functions and its derivatives with respect to material parameters is considered to solve the inverse problem (see [2] and [3]). However, this approach cannot model the reflection of the wave over interfaces.
- In the case of a thin layered solid, the wave equation is solved analytically using a double Laplace transform and the Cagniard method to process the reflection over the interface (see [4]). The limitation of this approach is that the solid domain has to be simple to be able to find the analytical solution. Thus, it is possible to recover only one thin layer.

In this paper, we propose to estimate the mechanical properties of a solid with multiple thin layers using a numerical solution of the wave equation. It is based on a finite element discretization which is relatively fast in computation (see [5] and [6]). Thus it is possible to consider a complex structure with multiple layers and run simulations thousands of times without the burden of demanding computations. The inverse problem is formulated as a nonlinear least squares problem where a cost function has to be minimized to get an estimation of the parameters.

Usually, this is done by computing the gradient of the cost function using the adjoint state method (see [7]). However, our mathematical model for the wave propagation in a layered solid is too complex to be used with the adjoint state method. So we propose to use instead a stochastic and global optimization technique, named Simultaneous Perturbation Stochastic Approximation (SPSA). We compare this method to Markov Chain Monte Carlo (MCMC) and show that SPSA is more interesting due to its cost and accuracy.

Finally, we study the stability of the solution of the inverse problem against random noise in the pressure measures.

2. THE FORWARD PROBLEM: A TWO-DIMENSIONAL NUMERICAL MODEL

2.1. The Physical Model

Our study starts with a two-dimensional mathematical model that consists in finding $(p, v_f, v_s$ and $\sigma)$ that verify the following system of partial differential equations and over an infinite domain constituted by a solid medium (Ω_s) and a fluid medium (Ω_f) and separated by an interface (Γ):

$$\begin{cases} \dfrac{\partial p}{\partial t} + c_f^2 \rho_f \ \mathrm{div}\ v_f & = 0 \ (\Omega_f) & (1.1) \\[2mm] \rho_f \dfrac{\partial v_f}{\partial t} + \nabla p & = 0 \ (\Omega_f) & (1.2) \\[2mm] A \dfrac{\partial \sigma}{\partial t} - \epsilon(v_s) & = 0 \ (\Omega_s) & (1.3) \\[2mm] \rho_s \dfrac{\partial v_s}{\partial t} - \ \mathrm{div}\ \sigma & = 0 \ (\Omega_s) & (1.4) \\[2mm] v_s.n & = v_f.n \ (\Gamma) & (1.5) \\[2mm] \sigma.n & = -p.n \ (\Gamma) & (1.6) \end{cases} \qquad (1)$$

where the unknowns are

- p: the pressure field in the fluid medium
- v_f: the velocity field in the fluid
- σ: the stress tensor field in the solid
- v_s: the velocity field in the solid

and the physical parameters are

- c_f: the wave propagation speed in the fluid
- ρ_f: the fluid density
- ρ_s: the solid density
- A: the inverse of the elasticity tensor which is function of λ and μ, the Lame's coefficients.

This model has been derived by [6]. Its main feature is that two media are clearly defined allowing us to consider independent discretization for each of the two domains (fluid and solid). These discretizations are based on the following variational formulations.

2.2. Numerical Model

Let:

- $L^2(\Omega) = \{w \mid \int_\Omega w^2 < \infty\}$ and
- $H_{\mathrm{div}}(\Omega) = \{w \in (L^2(\Omega))^2 \mid \mathrm{div}\ w \in L^2(\Omega)\}$,

be respectively the Hilbert spaces of square integrable functions in the sense of Lebesgue and the "H-div" space of vectorial functions which have each coordinate in $(L^2(\Omega))$ and its divergence also in $(L^2(\Omega))$

The semi-variational formulation for the fluid medium is obtained after multiplying Equations (1.1) and (1.2) by a pair of test functions $(q, w_f) \in ([0, T] \times L^2(\Omega_f)) \times ([0, T] \times H_{\text{div}}(\Omega_f))$, then integrating over Ω_f and using Green's formula, with boundary and interfaces conditions on Ω_f. Since this process involves only the space variables and leaves out the time variable, it leads to the following (semi-)variational formulation established by [6].

Find $(p, v_f) \in ([0, T] \times L^2(\Omega_f)) \times ([0, T] \times H_{\text{div}}(\Omega_f))$ such that:

$$\begin{cases} \dfrac{d}{dt} \int_{\Omega_f} \dfrac{pq}{c_f^2 \rho_f} + \int_{\Omega_f} q \ \text{div} \ v_f & = 0 \quad (2.1) \\[2ex] \dfrac{d}{dt} \int_{\Omega_f} \rho_f v_f . w_f - \int_{\Omega_f} p : \text{div} : w_f - \int_{\Gamma} p(w_f . n) & = 0 \quad (2.2) \end{cases} \qquad (2)$$

$\forall (q, w_f) \in ([0, T] \times L^2(\Omega_f)) \times ([0, T] \times H_{\text{div}}(\Omega_f))$.

Similarly for the solid domain, one finds the following (semi-)variational formulation for the solid.

Find $v_s \in [0, T] \times (L^2(\Omega_s))^2$ and $\sigma \in [0, T] \times H_{\text{div}}^{\text{sym}}(\Omega_s)$, $H_{\text{div}}^{\text{sym}}(\Omega_s) = \{\sigma | \sigma \in (L^2(\Omega_s))^4, \ \text{div} : \sigma_i \in (L^2(\Omega_s))^2 \ \text{and} \ \sigma_{ij} = \sigma_{ji}\}$:

$$\begin{cases} \dfrac{d}{dt} \int_{\Omega_s} A\sigma . \xi + \int_{\Omega_s} \text{div} \ \xi . v_s - \int_{\Gamma} (\xi n) . vs & = 0 \quad (3.1) \\[2ex] \dfrac{d}{dt} \int_{\Omega_s} \rho_s v_s . w_s - \int_{\Omega_s} w_s . \text{div} \ \sigma & = 0 \quad (3.2) \end{cases} \qquad (3)$$

$\forall (w_s, \xi) \in (([0, T] \times (L^2(\Omega_s))^2) \times ([0, T] \times H_{\text{div}}^{\text{sym}}))$.

The finite elements used to discretize (1) are mixed finite elements and are taken from [6]. The pressure $p \in L^2(\Omega_f)$ is approximated by a finite-element piecewise constant function in $Q_0(\Omega_f) \subset L^2(\Omega_f)$; for $v_f \in H_{\text{div}}(\Omega_f)$: we use the Raviart–Thomas finite elements

$$RT_0(H)\{q \quad \in \quad L^2(\Omega_f) : \forall e \in E, \exists a \in \mathbf{R}^2 \text{ and } b \in \mathbf{R} \mid q(x) = ax + b$$

$$\text{and } \forall e \quad \in \quad E \ [q]_e . v_e = 0\} \qquad (4)$$

where H and E denote the sets of meshes and edges, $[q]_e$ is the jump of q over an edge e and v_e is a normal vector to e.

For σ and v_s we used the finite elements proposed by [5], with regular rectangular meshes, which present the following advantages:

- the mass matrices are block diagonal,
- the symmetry of the stress tensor does not need to be imposed. Such condition becomes "natural" rather than "essential".

Note however that, in such method, one disadvantage is that curved interfaces would be approximated with a staircase discretization.

The finite elements chosen to approach the stress tensor are those described in [5]. This will be considered only for the case of a horizontal interface between the solid and the fluid. The degrees of freedom of the stresses are all associated to the nodes of the elements. In order to conform to the transmission conditions at the interface, one imposes that $\sigma_{12} = 0$ for all nodes at the fluid–solid interface.

Applying the standard finite element procedure that consists in seeking $P \in Q_0(\Omega_f)$, $V_f \in RT_0$, $V_s \in Q_0(\Omega_s)$ and $\Sigma \in Q_{\mathrm{div}}^1 - Q^0$ that verify the semi-variational formulation (3) and (2), P, V_f, V_s and Σ are evaluated respectively at instants $n, n + 1/2, n, n + 1/2$:

$$
\begin{cases}
M_p \dfrac{P^{n+\frac{1}{2}} - P^{n-\frac{1}{2}}}{\Delta t} + D_f V_f^n & = 0 \quad (5.1) \\[2ex]
M_f \dfrac{V_f^{n+1} - V_f^n}{\Delta t} - D_f^t P^{n+\frac{1}{2}} + B \dfrac{\Sigma^{n+1} + \Sigma^n}{2} & = 0 \quad (5.2) \\[2ex]
M_\Sigma \dfrac{\Sigma^{n+1} - \Sigma^n}{\Delta t} + D_s^t V_s^{n+\frac{1}{2}} - B^t \dfrac{V_f^{n+1} + V_f^n}{2} & = 0 \quad (5.3) \\[2ex]
M_s \dfrac{V_s^{n+\frac{1}{2}} - V_s^{n-\frac{1}{2}}}{\Delta t} + D_s \Sigma^n & = 0 \quad (5.4).
\end{cases}
\tag{5}
$$

This computational model provides a solution of the forward problem given the set of physical parameters for the fluid–solid pair, $\theta = (\rho, \lambda, \mu)$. In such model, the matrices D_f and D_s represent discrete divergence operators and B is the discrete trace operator. Since p and v_s are piecewise constant, M_p and M_s are diagonal matrices. The mass-lumping technique described in [5] is applied in order to obtain a block-diagonal matrix $M_\Sigma (5 \times 5)$ and a diagonal matrix M_f.

3. THE BAYESIAN MODEL AND THE INVERSE PROBLEM

The inverse problem starts with experiments consisting in placing receivers at various points of the fluid, then stimulating the pair solid–fluid by sending seismic waves from the fluid. The receivers measure the variation of the pressure resulting from such stimulation. An inverse problem formulation would consist in finding the vector of parameters $\theta = (\rho_l, \lambda_l, \mu_l)$ $l = 1, \ldots, n_l$ from the pressure measurements, n_l is the number of layers of the solid. It aims at computing estimations of θ on the basis of a given set of data $y := \{y_{ij} = y(x_i, t_j)\}$ representing noisy measures of the pressure at a set of points $\{x_i\}$ in the fluid, at times $\{t_j\}$. The formulation of the inverse problem considered in the paper is a Bayesian formulation. It relates pressure measures to mechanical properties in a probabilistic sense and using the forward problem.

Specifically, let

$$
y = P(\theta) + \epsilon
\tag{6}
$$

be the input data of the inverse problem, where $\epsilon = \{\epsilon_{ij}\}$ is a blank noise following a Gaussian law $N(0, s^2)$; s^2 is the fixed variance of the noise and where $P(\theta)$ is the pressure computed by the forward model with parameter θ.

Estimation of θ requires $p(\theta|y)$ the probability distribution of θ given y. The Bayesian formula reads

$$p(\theta|y) = \frac{1}{p(y)} p(y|\theta) p(\theta). \tag{7}$$

Use of (7) is coupled with the following assumptions:

- The prior probability density $(p(\theta))$ is a knowledge on θ before making the measurements. It is the range of acceptable values for the physical parameters such that the solid having these values of parameters can be found in nature. It is supposed that the value of each parameter fall into the domain of acceptable values $D_\theta = \prod_i [\theta_{i,\min}, \theta_{i,\max}]$ with a uniform probability. The prior probability density $(p(\theta))$ is given by

$$p(\theta) \propto \begin{cases} 1 & \text{if } \theta \in D_\theta \\ 0 & \text{elsewhere.} \end{cases} \tag{8}$$

- The likelihood probability density $(p(y|\theta))$ is the relationship which maps the physical parameters to the solution of the forward problem and computed with numerical simulations. In Equation (6) $P(\theta, x, t)$ is not a random variable so the expression of the likelihood probability distribution $p(y|\theta)$ can be deduced from the one of ε which is a Gaussian distribution of mean zero and variance s^2

$$p(y|\theta) \propto \exp\left(\frac{-1}{2} \|y - P(\theta)\|_F^2 \right) \tag{9}$$

   where $\|.\|_F^2$ is the Frobenius norm.
- Equation (7) with (8) and (9) leads to the following expression of the posterior probability density:

$$p(\theta|y) \propto \begin{cases} \exp\left(\frac{-1}{2} \|y - P(\theta)\|_F^2 \right) & \text{if } \theta \in D_\theta \\ 0 & \text{elsewhere.} \end{cases} \tag{10}$$

The normalization constant $\left(\frac{1}{p(y)} \right)$ is a constant with respect to θ that verifies

$$\int_{D_\theta} \frac{1}{p(y)} p(y|\theta) p(\theta) = 1. \tag{11}$$

This constant is not calculated and the expression of the posterior probability density is unnormalized. In (10), the pressure $P(\theta)$ is the result of a forward simulation (5), with ρ_s

and A defined by θ. Thus it must be kept in mind that finding the value of $p(\theta|y)$ for a certain θ is costly and an algorithm which requires a few number of evaluation of $p(\theta|y)$ has to be used.

On this basis, we implement the following estimators:

1. The first seeks

$$\Theta(y) = E(\theta|y) = \int \theta\ p(\theta|y)\ d\theta,$$

   the expectation of θ given y.
   This estimator is evaluated using the MCMC algorithm which is described in Section (4). Results are shown for an accelerated form of this method.
2. The second computes the maximum of the posterior probability $p(\theta|y)$:

$$\theta^* = \arg \max_{\theta \in D_\theta}\ p(\theta|y), \tag{12}$$

   where arg provides the value of θ where the given objective function ($p(\theta|y)$) is maximal. This second estimator is computed with the SPSA method which is detailed in Section 5. This maximization problem is equivalent to minimizing $-\log p(\theta|y) = ||y - P(\theta)||_F^2$. So in our case, this estimation of the parameters is also the solution of the least squares error problem:

$$\theta^* = \arg \min_{\theta \in D_\theta}\ ||y - P(\theta)||_F^2. \tag{13}$$

4. MARKOV CHAIN MONTE CARLO

The first estimation of the mechanical properties of interest is its expected value:

$$E(\theta|y) = \int \theta\ p(\theta|y)\ d\theta.$$

Monte Carlo methods were first developed to estimate integrals that cannot be evaluated analytically (see [8]). Although the term "Monte Carlo methods" includes several statistical techniques, this paper is based on a standard Monte Carlo type integration. In such a method, approximation of the integral is done through draws of N dependent, identically distributed samples of θ, say θ_k, $k = 1, \ldots, N$ from the probability density $p(\theta|y)$. The average of these samples is then computed, leading to the estimate:

$$\int \theta\ p(\theta|y)\ d\theta - \frac{1}{N} \sum_{k=1}^{N} \theta_k = O(1/\sqrt{N}), \quad \theta_k \sim p(\theta|y) \tag{14}$$

The Metropolis–Hastings generates a sequence of samples (θ_k) which form a Markov chain with the limiting distribution $p(\theta|y)$. This method uses a proposal density $q(.|\theta)$. Its usage is shown next in the description of the algorithm. A common choice for this distribution is the random walk with a uniform probability. Specifically, let $D_{\theta_k,\delta}$ be the parallelepiped of center θ_k and edges 2δ

$$D_{\theta_k,\delta} = \prod_i [\theta_{k,i} - \delta_i, \theta_{k,i} + \delta_i].$$

Let also $U(D_{\theta_k,\delta})$ be the uniform distribution over $D_{\theta_k,\delta}$ and null outside $D_{\theta_k,\delta}$. We define

$$q(.|\theta_k) = U(D_{\theta_k,\delta}).$$

Note that $q(C|\theta_k) = q(\theta_k|C)$. The standard M-H algorithm includes two steps and is described as follows:

1. **Generation step:** from a state θ_k, a candidate C is proposed by drawing at random on the basis of $q(.|\theta_k)$. Simultaneously, a number r is randomly drawn from the interval $(0,1)$.
2. **Acceptance step:** The probability

$$\alpha(C, \theta_k) = \min\left(\frac{p(C|y)}{p(\theta_k|y)}, 1\right) \tag{15}$$

   is evaluated and the candidate C is accepted if $\alpha(C, \theta) > r$. Thus, $\theta_{k+1} = C$ if the candidate is accepted and otherwise, $\theta_{k+1} = \theta_k$, i.e. the Markov chain remains at the same state.

This Markov chain generation is highly expensive because it requires many evaluations of $p(C|y)$ by running a numerical simulation. To overcome this difficulty, we use an accelerated MCMC based on the algorithm in ([9]).

4.1. An Accelerated Version of Markov chain Monte Carlo

The main idea for reducing the cost of computations is to run simulations for the candidates that will likely be accepted. Thus we avoid computations for the candidates that are likely to be refused. To guess whether a candidate C will be accepted or not $p(C|y)$ is evaluated approximately. Unlike in ([9]) we use a linear interpolation, named $p^*(.|y)$. Thus if the predicted probability is high, a simulation is run to evaluate exactly $p(C|y)$.

   The modification (acceleration) of the standard algorithm is in the insertion of an intermediate step between the generation and the acceptance steps of the standard M-H:

1. At θ_k generate a proposal C from $q(\cdot|\theta_k)$.
2. A number r is randomly drawn from the interval $(0,1)$. Evaluate

$$\alpha_{\text{pred}}(C, \theta_k) = \min \left(\frac{p^*(C|y)}{p^*(\theta_k|y)}, 1 \right) \tag{16}$$

C is promoted to be a candidate to the standard M-H algorithm if $\alpha_{\text{pred}}(C, \theta) > r$. Otherwise, pose $\theta_{k+1} = \theta_k$.

3. Again a number r is randomly drawn from the interval (0,1). Evaluate

$$\alpha(C, \theta_k) = \min \left(\frac{p(C|y)}{p(\theta_k|y)}, 1 \right) \tag{17}$$

accept $\theta_{k+1} = C$ if $\alpha(C, \theta) > r$; otherwise reject C, $\theta_{k+1} = \theta_k$.

So we recall here that the estimation of the mechanical properties θ is done by generating a sequence of samples $\theta_k \sim p(\theta|y)$, $k = 1, \ldots, N$ from the accelerated version of MCMC and by evaluating its average value

$$E(\theta|y) \approx \frac{1}{N} \sum_{k=1}^{N} \theta_k. \tag{18}$$

In order to get a reliable result with MCMC, the proposal distribution $q(C|\theta_k)$ has to be tuned very carefully because MCMC is very sensitive to its proposal distribution. Several tests have to be done to be sure of the reliability of the results. These tests concern the accuracy, the convergence and the uniqueness of the solution. They are detailed in Section 7.1.1.

5. SIMULTANEOUS PERTURBATION STOCHASTIC APPROXIMATION ALGORITHM

The second estimator is the maximum a posteriori which is the solution of the following minimization problem:

$$\theta^* = \arg \min_{\theta \in D_\theta} L(\theta) \tag{19}$$

where $L(\theta) = ||y - P(\theta, x_i, t_j)||_F^2$.

We recall that $\theta = \{(\rho_l, \lambda_l \mu_l), \ l = 1, \ldots, n_l\}$ is the set of mechanical properties to find with n_l the number of layers in the solid. The loss function $L(\theta)$ represents the misfit function between the input pressure measures and the simulated pressure corresponding to θ. In what follows, it is assumed that $L(\theta)$ is a scalar differentiable function of the $3n_l$-dimensional vector θ representing the unknown parameters and the estimator θ^* should verify

$$\frac{\partial L}{\partial \theta}(\theta^*) = 0.$$

Given that the exact value of the gradient L with respect to θ is not available, the gradient of L is computed using a stochastic approximation based on only two evaluations of L.

The SPSA algorithm is of the form

$$\theta_{k+1} = \Psi(\theta_k - a_k g(\theta_k)), \tag{20}$$

where Ψ is a projection application that keeps $\theta_k \in D_\theta$ and $g(\theta_k)$ is a stochastic estimation of the true gradient $\partial L/\partial \theta$.

In the SPSA method, one uses a special approximation for the gradient that requires only two evaluations of $L(\theta)$. All the variables are perturbed at the same time by drawing two random points $\theta_k + c_k \Delta_k$ and $\theta_k - c_k \Delta_k$ centered at θ_k to form the gradient approximation:

$$g(\theta_k) = \frac{L(\theta_k + c_k \Delta_k) - L(\theta_k - c_k \Delta_k)}{2c_k} (\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \ldots, \Delta_{3n_l}^{-1})^T \tag{21}$$

Δ_k is the perturbation random variable vector of mean zero and c_k is some positive number. In [10, 21] it is shown that this stochastic approximation of the gradient makes SPSA a global optimizer under certain conditions.

The SPSA algorithm is summarized in the following steps:

1. **Initialization step:** With the SPSA, the sequences a_k and c_k are of the form $a_k = a/(A + k)^\alpha$ and $c_k = c/k^\gamma$. Set counter index $k = 1$. Define values for the positive coefficients a, c, A, α, γ, and the maximal number of iterations $k_{\max}$.
2. **Generation of the simultaneous perturbation vector:** Generate a $3n_l$-dimensional random perturbation vector Δ_k, where each component of Δ_k is independently generated from a zero-mean probability distribution satisfying the previous conditions. A simple (and theoretically valid) choice for each component of Δ_k is to use a Bernoulli distribution.
3. **Loss function evaluations:** Compute $L(\theta_k + c_k \Delta_k)$ and $L(\theta_k - c_k \Delta_k)$ which involves numerical simulation of the forward problem.
4. **Gradient approximation and updating θ_k:** Generate the approximate gradient $g(\theta_k)$ from (21) and update θ_k using Equation (20).
5. **Iteration or termination:** Return to step 2 and increment k by 1 or terminate the algorithm if the maximum allowable number of iterations is reached.

The choice of the gain sequences (a_k and c_k) is critical to the performance of SPSA. In [11], it is provided some guidance on picking these coefficients in a practically effective manner. In [12] and [13], it is presented sufficient conditions for convergence of the SPSA iterates.

5.1. Confidence Interval for θ^*

We suppose that the gain sequences have the standard form $a_k = a/(k + 1 + A)^\alpha$ and $c_k = c/(k + 1)^\gamma$, with a, c, α, γ strictly positive, $A \geq 0$, $\beta = \alpha - 2\gamma > 0$ and $3\gamma - 2\alpha \geq 0$. In [14] the convergence of the method is established in a stochastic meaning:

$$k^{\beta/2}(\theta_k - \theta^*)\xrightarrow{\text{dist}} N(\mu, \Sigma) \ \ as \ k \to \infty.$$

The mean μ depends on both the Hessian matrix and the third derivative of $L(\theta^*)$ and Σ depends on the Hessian at $L(\theta^*)$ ([14]). If we take $3\gamma - 2\alpha > 0$, $\mu = 0$. We propose a confidence interval for θ^* by using the standard definition of the mean value and the covariance for the random variable $k^{\beta/2}(\theta_k - \theta^*)$ (for a sequence θ_k, $k = 1, \ldots, N$, sufficiently long):

1. The mean value of the random variable $k^{\beta/2}(\theta_k - \theta^*)$ is computed as follows:

$$\mu = \frac{1}{N}\sum_{k=1}^{N} k^{\beta/2}(\theta_k - \theta^*) = 0. \tag{22}$$

2. We deduce the value of θ^*

$$\theta^* = \frac{\sum_{k=1}^{N} k^{\beta/2}\theta_k}{\sum_{k=1}^{N} k^{\beta/2}}. \tag{23}$$

3. The covariance of the estimation of θ^* is

$$\Sigma = \frac{1}{N}\sum_{k=1}^{N} k^{\beta}(\theta_k - \theta^*)(\theta_k - \theta^*)^T. \tag{24}$$

Thus the 95% confidence interval for each parameter θ_i is equal to $\theta_i^* \pm 2\sqrt{\Sigma_{ii}}$.

6. SENSITIVITY ANALYSIS

The purpose of the sensitivity analysis is to find the threshold of noise in the pressure measures above which some parameters are not identifiable. We assume that, for any $\theta \in \mathbf{R}^p$ ($p = 3n_l$), the forward problem has a unique solution. We denote by $P(\theta) \in \mathbf{R}^m$ the pressure $P(\theta, x_i, t_j)$ at points x_i and instants t_j, with $m \geq p$. We assume that $P(0) = 0$ and that $\frac{\partial P}{\partial \theta} \in \mathbf{R}^{m \times p}$ exists and is of full rank.

Let $y = P(\theta) + \varepsilon = P(\theta + \delta\theta)$ be the measured pressure. We compute a bound for the error $\delta\theta$ in the estimated parameter using a first-order analysis. We write a linear approximation of y and $P(0) = 0$ around θ (see [15]):

$$\begin{cases} \varepsilon \approx \dfrac{\partial P}{\partial \theta}(\theta)\delta\theta \\[2mm] P(\theta) \approx \dfrac{\partial P}{\partial \theta}(\theta)\theta. \end{cases} \tag{25}$$
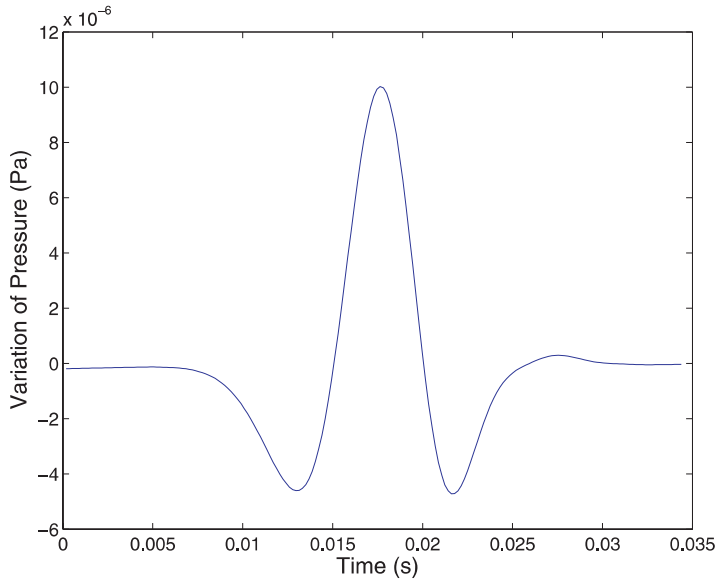
Figure 3. Forward numerical model: time variation of the pressure computed at a point in the fluid.

Let $\sigma_{\min}$ and $\sigma_{\max}$ the smallest non-zero and the largest singular values of $\frac{\partial P}{\partial \theta}$; then

$$\begin{cases} ||\delta\theta|| \leq \dfrac{1}{\sigma_{\min}}||\varepsilon|| \\ ||P(\theta)|| \leq \sigma_{\max}||\theta||. \end{cases} \tag{26}$$

Thus

$$\frac{||\delta\theta||}{||\theta||} \leq K\frac{||\varepsilon||}{||P||} \tag{27}$$

where $K = \dfrac{\sigma_{\max}}{\sigma_{\min}}$ is the condition number of $\frac{\partial P}{\partial \theta}$. Therefore, at first order, the noise is amplified by the condition number of the forward problem.

7. NUMERICAL RESULTS

The computational software that does the numerical simulation of the wave propagation was developed by team ONDES at INRIA Rocquencourt [5, 6]. It is considered to be fast, since at each time step, computing the solution at each node requires a local scheme, unlike other existing codes (see [5]) that use a global system to advance the solution in time. Figure 3

Table 1. Estimation of the density and the Lamé's coefficients of a homogeneous solid with 19 000 samples from the accelerated Markov chain Monte Carlo (6000 simulations) and level of noise $\leq 6\%$.

Physical Parameters	Exact Value (SI)	Confidence Interval	Error between estimated and true parameters
λ	11.5×10^9	$10.9 \times 10^9 \pm 2.6\%$	5.2%
μ	6×10^9	$6.5 \times 10^9 \pm 2\%$	8%
ρ	1850	$1867 \pm 0.15\%$	0.9%

shows the time variation of the pressure at a given point of the fluid. The input data of the inverse problem are pressure measures obtained synthetically from this software.

7.1. Test Case with a Homogeneous Solid

We compare accelerated MCMC and SPSA in the case of a homogeneous solid medium where only three parameters (λ, μ, ρ_s) are to be estimated. The source is placed at 50 m above the solid–fluid interface. A standard choice for the transmitted signal is the derivative of a Gaussian. Its frequency is 100 Hz and its amplitude is 1000 Pa. Twenty-six receivers are placed at the same depth of the source and evenly spaced by $5m$. Seventy-eight pressure measures are considered at each receiver. To simulate an observational measure error, a Gaussian noise $\sim N(0, s^2)$ is added to the pressure computed by the forward problem. The covariance s of the noise is taken equal to 6% of Frobenius norm of the measures array.

7.1.1.   Results with MCMC

We estimate the three parameters of the solid using the accelerated MCMC algorithm. The main results are shown in Table 1.

   The validity of the results with MCMC is checked through four distinct tests:

1. *Convergence of our estimate to $E(\theta|y)$:* We want to know if the number of samples is sufficient to have a correct estimation of our parameters. The plot of the average $E(\theta|y) = \frac{1}{N}\sum_{k=1}^{N}\theta_k$ with respect to the first n samples of the chain is shown Figure 4. One must see that this average becomes almost for $n > 15\,000$ samples so the convergence is reached (see [16]).

2. *Sampling from $p(\theta|y)$:* We need to check if the samples follow $p(\theta|y)$, the limiting distribution of the Markov chain. One must expect the same evaluation for the parameters at the convergence with two different initial points (see [16]). Two Markov chains of length 19 000 samples are run, one starting with $\theta_{\min}$ the smallest acceptable values of θ and the other starting with $\theta_{\max}$ the largest acceptable values of θ. The results are compared in Table 2. The two results are not the same but they are close enough to be able to consider that a length of 19 000 samples for the Markov chain is sufficient and the samples are quite distributed according to $p(C|y)$.
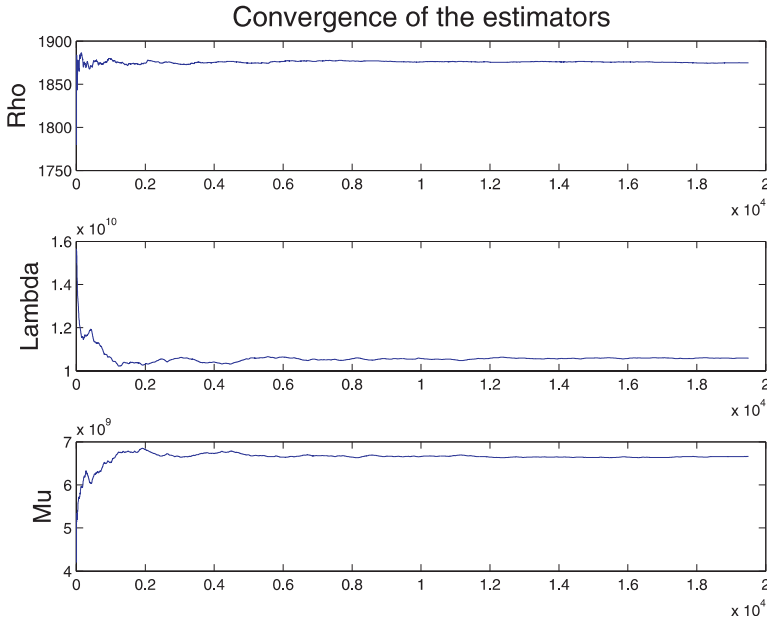
Figure 4. Convergence of Markov chain Monte Carlo: estimation of the parameters with respect to the size of the Markov chain.

Table 2. Convergence of Markov chain Monte Carlo (MCMC) (test 3): Two estimations of the density and the Lamé coefficients with two different starting points and with 19 000 samples from the accelerated MCMC (6000 simulations) and level of noise $\leq 6\%$.

Physical Parameters	Confidence Interval with $\theta_{\min}$	Confidence Interval with $\theta_{\max}$
λ	$11.1 \times 10^9 \pm 2.8\%$	$10.8 \times 10^9 \pm 2.7\%$
μ	$5.82 \times 10^9 \pm 2.5\%$	$6.14 \times 10^9 \pm 2.6\%$
ρ	$1827 \pm 0.21\%$	$1911 \pm 0.4\%$

3. *Uniqueness of the solution:* the plot of the frequency of the samples generated by the chain is shown in Figure 5 from which one can verify that the posterior probability has one mode which means that the inverse problem has a unique solution.

4. *Variance of our estimation:* the correlation between the samples separated by s $C(s)' = \text{cov}(\theta_k, \theta_{k+s})$ gives a diagnosis on the behavior of the Markov chain. If the sequence θ_k are sampled from the desired probability $p(\theta|y)$, one must find a constant M for which $C(s)$ tends to zero for $s > M$ (see [17, 18]). Figure 6 shows the correlation function $C(s)$ which verifies the previous requirement. From this figure it is possible to compute the variance of the estimation of the parameters obtained by MCMC $\text{var}(\overline{\theta}_{\text{MC}})$. It is given by the formula $\text{var}(\overline{\theta}_{\text{MC}}) = \tau \frac{\text{var}(\theta)}{n}$, where $\text{var}(\theta)$ is the variance of the samples θ_k but supposed independent, and τ is the integrated autocovariance time given by $\overline{\tau} = 1 + 2 \sum_{s=1}^{M} C(s)/C(0)$ (see [17, 18]).
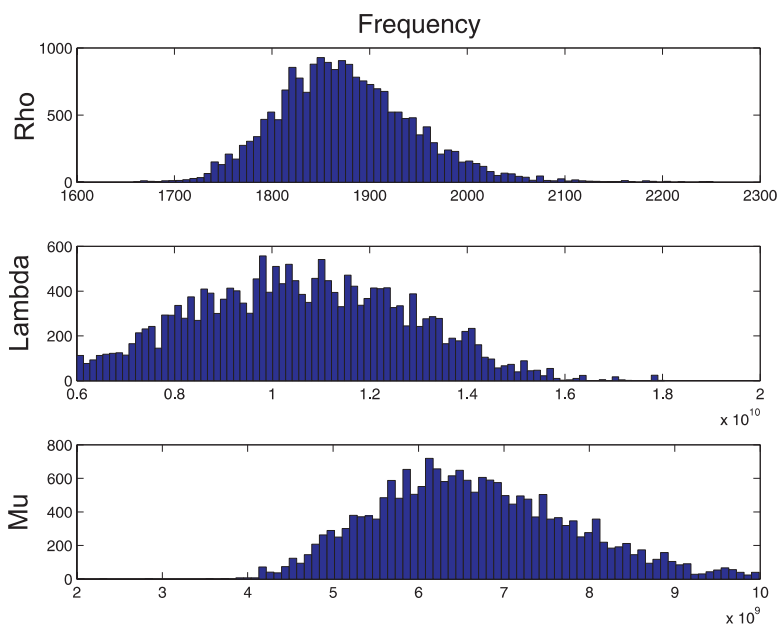
Figure 5. Verification of the uniqueness of the solution of Markov chain Monte Carlo: mode of the posterior distribution.
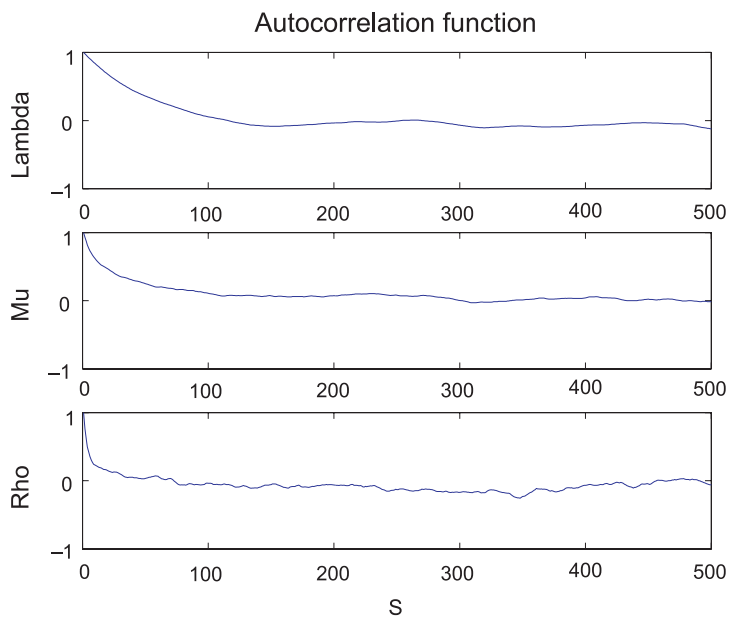

Figure 6. Sampling from $p\theta|y$: autocorrelation function of each parameter.
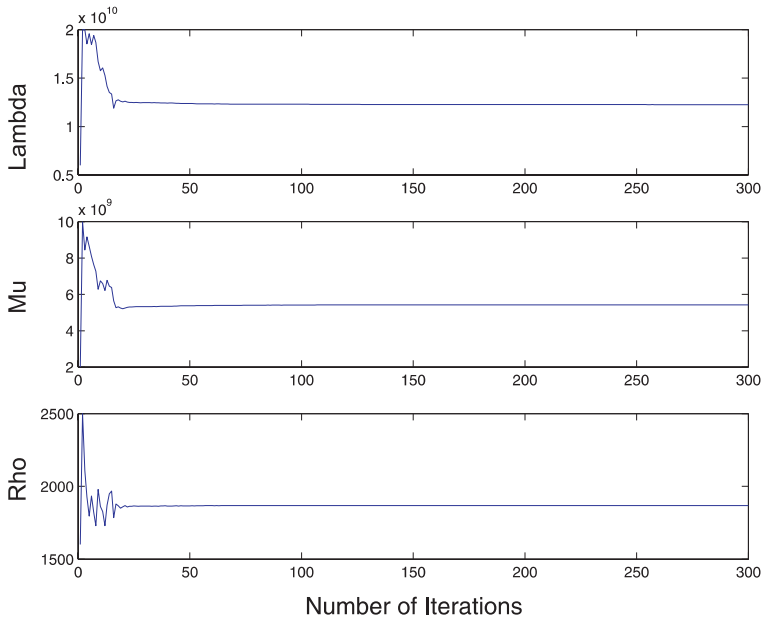
Figure 7. Convergence of simultaneous perturbation stochastic approximation: trajectories of the three mechanical parameters with respect to the number of iterations in the case of a one-layer solid.

Table 3. Estimation of the density and the Lamé coefficients of a homogeneous solid with SPSA and level of noise $\leq 6\%$ (700 iterations).

θ	Exact Values (SI)	Confidence Intervals	Errors
λ	11.5×10^9	$12.2 \times 10^9 \pm 6.12\%$	6.6%
μ	6×10^9	$5.4 \times 10^9 \pm 7.2\%$	9.5%
ρ	1850	$1868 \pm 0.83\%$	1%

In the case of a homogeneous solid, MCMC gives results with an acceptable accuracy but it is highly expensive in computation. Moreover, it is difficult to carry out results with MCMC because it requires tuning the parameters of the algorithm carefully to obtain reliable results. Therefore we compare MCMC with the algorithm SPSA, which is easier to implement.

7.1.2. Results with simultaneous perturbation stochastic approximation

With SPSA we take the same pressure measures and level of noise as with MCMC. The estimation with this algorithm is shown in Table 3. The trajectories of parameters are in Figure 7.

We see that SPSA is much less expensive than accelerated MCMC (700 simulations instead of 6000) and gives the same accuracy. Thus we consider SPSA with a multi-layer solid.

Table 4. Estimation of the density and the Lamé coefficients of the three layers with simultaneous perturbation stochastic approximation and level of noise $= 0\%, 1\%$ and 3.5% (6000 iterations).

θ	Exact Values (SI)	Estimated values with 0% noise (estimation error)	Estimated values with 1% noise (estimation error)	Estimated values with 3.5% noise (estimation error)
λ_1	11.5	11.59 (0.77%)	12.65 (9.09%)	39.15 (70.62%)
μ_1	6	5.97 (0.50%)	5.92 (1.35%)	5.10 (17.64%)
ρ_1	1700	1697 (0.17%)	1708 (0.46%)	1665 (2.10%)
λ_2	9	9.09 (0.99%)	8.49 (6.00%)	10.05 (10.44%)
μ_2	7	6.96 (0.57%)	7.10 (1.40%)	6.41 (9.20%)
ρ_2	2000	1993 (0.35%)	2052 (2.53%)	1972 (1.41%)
λ_3	11.5	11.65 (1.28%)	11.22 (2.49%)	16.18 (28.92%)
μ_3	6	5.96 (0.67%)	6.43 (6.68%)	6.32 (5.06%)
ρ_3	2400	2408 (0.33%)	2500 (4%)	2500 (4%)

7.2. Test Case with Three-Layer Solid

The numerical experiment in this case is very similar to the previous one. The source is placed at 50 m above the solid–fluid interface and each layer has a thickness of $5m$. A standard choice for the transmitted signal is the derivative of a Gaussian. Its frequency is 100 Hz and its amplitude is 1000 Pa. Twenty-six receivers are placed at the same depth of the source and evenly spaced by 5 m. 208 pressure measures are considered at each receiver. To simulate an observational measure error, a Gaussian noise $\sim N(0, s^2)$ is added to the pressure measures. We consider three levels of noise $s = 0\%$ (no noise), $s = 1\%$ and $s = 3.5\%$ of Frobenius norm of the measures array.

The layers are considered thin because the wavelength of the source which is related to the frequency is less than the thickness of the layers. This configuration of the solid makes the reflected signals overlap and the inverse problem difficult to solve.

Before solving this case, we make a sensitivity analysis to get the threshold of error in the pressure measures above which the parameters are not identifiable. We compute $\frac{\partial P}{\partial \theta}$ by using a finite difference approximation and an approximate condition number of $\frac{\partial P}{\partial \theta}$. We obtain $K \approx 28$. Therefore, following (27) and in order to get $\frac{||\delta\theta||}{||\theta||} \leq 1$, $\frac{||\varepsilon||}{||P||} \leq K^{-1} = 0.035$ (see [19] and [20]). This motivates our choice of s above.

Results are summarized in table 4 and the trajectories of the parameters are depicted in Figure 8. We see that SPSA gives good results with pressure measures without noise, and with noise equal to 1%. But when the noise is equal to the theoretical threshold given by the sensitivity analysis, some parameters are badly estimated. These results are expected due to the high sensitivity of the inverse problem to the noise in the pressure measures.
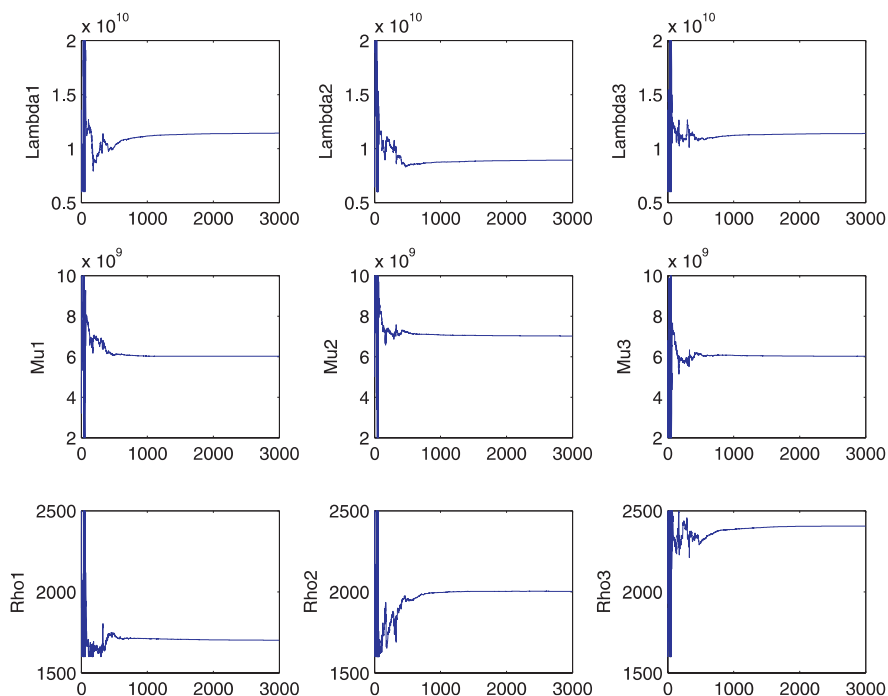
Figure 8. Convergence of simultaneous perturbation stochastic approximation: trajectories of the nine parameters with respect to the number of iterations in the case of a three layers solid, noise $= 1\%$.

8. CONCLUDING REMARKS

We are interested in the inverse problem which is the estimation of the density and the Lamé coefficients of a layered solid in contact with a fluid medium from the variation of the pressure in the fluid due to the propagation of a seismic wave. The layers are considered thin because their thickness is less than the wavelength.

The inversion is based on synthetic pressure measures and numerical simulations of the elastic wave propagation. Two stochastic methods are implemented and compared with respect to their computational cost and accuracy.

The first method, MCMC, gives for this problem an estimation of the parameters of interest with an acceptable error but its very costly in computation so only the case of one-layer solid has been tested.

The second method, SPSA, gives the same level of accuracy as MCMC when compared for a one-layer solid and it has the advantage of being much less expensive in computation. It is possible with this method to go further and solve the case of a three-layer solid with about the same cost needed with MCMC for a one-layer solid.

REFERENCES

[1] Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A., and Wyld, H. W., Tomographic determination of velocity and depth in laterally varying media. *Geophysics,* 50, 903–923 (1985).

[2] Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation,* SIAM, Philadelphia, PA, 2005

[3] Cerveny, V. *Seismic Ray Theory*, Cambridge University Press, Cambridge, 2001.

[4] Ma, C.-C., Liu, S.-W., and Chang, C.-M. Inverse calculation of material parameters for a thin-layer system using transient elastic waves. *Acoustical Society of America Journal,* 112, 811 (2002)

[5] Tsogka, C. Modélisation Mathématique et Numérique de la Propagation des Ondes Élastiques Tridimensionnelles dans les Milieux Fissurés. *Ph.D thesis,* Chapters 1 and 2, Université Paris Dauphine-Paris IX, 1999.

[6] Diaz, J. and Joly, P. Robust High Order Non Conforming Finite Element Formulation for Time Domain Fluid-Structure. *Journal of Computational Acoustics,* 13, 403–431 (2005).

[7] Plessix, R.-E. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International,* 167, 495–503 (2006)

[8] Niederreiter, H. *Random Number Generation and Quasi-Monte Carlo Methods,* CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 63, vi, p.241, Society of Industrial Applied Mathematics, Philadelphia, PA, 1992.

[9] Andres Christen, J. and Fox, C. MCMC Using Approximation. *Journal of Computational and Graphical Statistics,* 14, 795–810 (2005).

[10] Maryak, J. L. and Chin, D. C. Global Random Optimization by Simultaneous Perturbation Stochastic Approximation, in *Proceedings of the American Control Conference, Arlington, VA,* pp. 756–762, 2001.

[11] Spall, J. C. Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization. *IEEE Transactions of Aerospace and Electronic Systems,* 34, 817–823 (1998).

[12] Spall, J. C. A stochastic approximation algorithm for large-dimensional systems in the Kiefer–Wolfowitz setting, in *Proceedings of IEEE Conference on Decision and Control*, pp. 1544–1548, 1988.

[13] Spall, J. C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automated Control*, 37, 332–341 (1992).

[14] Spall, J. C. *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*, Chapter 7, Wiley, Hoboken, NJ, 2003.

[15] Bjorck, A. *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, 1996.

[16] Robert, C. *Méthodes de Monte Carlo par chaines de Markov*, Economica, Paris, 1996.

[17] Meyer, S., Christensen, N. and Nicholls, G. Fast Bayesian Reconstruction of Chaotic Dynamical Systems Via Extended Kalman Filtering. *Physical Review*, 3535–3542 (2001).

[18] Tan, M.S., Fox, C. and Nicholls, G. *Physics 707SC Course Materials*, Chapters 7–9, Auckland University.

[19] Assous, F. and Collino, F. A numerical method for the exploration of sensitivity: the case of the identification of the 2D stratified elastic medium. *Inverse Problems*, 6, 487–513 (1990)

[20] Al Khoury, P., Chavant, G., Clément, F. and Hervé, P. Inversion of spectroscopic data, application on CO2 radiation of flame combustion. *Inverse Problems in Science and Engineering*, 13, 219–240 (2005)

[21] Gerecér, L. and Vágó, Z. The Mathematics of Noise-Free SPSA. *Proceedings of the IEEE Conference on Decision and Control*, pp. 4400–4405 (2001).