JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 96, NO. B5, PAGES 8057–8067, MAY 10, 1991

Electromagnetic Imaging and Simulated Annealing

DOMINIQUE GIBERT

Ecole Nationale Supérieure des Mines de Paris

JEAN VIRIEUX

Institut de Géodynamique, CNRS, Valbonne, France

In contrast with acoustical imaging methods, for which the wave field is dominated by propagation effects, electromagnetic imaging of conductive media suffers from the diffusive behavior of the electromagnetic field. An important question to address when working toward the achievement of electromagnetic imaging concerns the possibility of resolving the diffusion damping. Exact inversion will be looking at the solvability of the integral equation relating a diffusive field to its dual wavefield. This equation is ill posed because its Laplace-like kernel makes the inverse problem of finding the dual wave field a notoriously difficult (both numerically and mathematically) one. Stochastic inversion is another alternative based on least squares fitting. In this inverse problem approach, extracting the wave field is still a relatively instable process, although the L^2 misfit function for data without noise presents a global minimum. The simulated annealing overcomes this instability for parameterization of this problem designed as follows. The unknown wave field is expected to be a sequence of impulsive functions. The number of impulsive functions can be determined by using a statistical criterion, called AIC, which comes from the Prony technique. The simulated annealing is applied to the positions of the reflections, while the amplitudes, which are not taken as parameters, are obtained by linear fitting. The simulated annealing method proves to be efficient even in the presence of noise. Furthermore, this nonlinear numerical inversion furnishes statistical quantities which allows an estimation of the resolution. Simple synthetic examples illustrate the performance of the inversion, while a synthetic finite element example shows the final pseudo-seismic section to be processed by standard seismic migration techniques.

INTRODUCTION

For typical Earth conductivities the electromagnetic response is mainly controlled by a diffusion process broadening the signal and making it quite different from a seismic signal. Obtaining the electromagnetic response of an arbitrary conductive medium is in itself an already difficult task investigated with approaches as different as integral equations, in both time domain [*SanFilipo and Hohmann*, 1985] and frequency domain [*Wannamaker et al.*, 1984], or as finite difference formulations [*Oristaglio and Hohmann*, 1984]. Recovering conductive information is an even more difficult operation investigated with limited success [*Weidelt*, 1972, 1975; *Tripp et al.*, 1984; *Barnett*, 1984; *Smith*, 1988; *Tarits*, 1989].

Although a mathematical parallelism between electromagnetic and seismic signals in layered media has been underlined previously [*Weidelt*, 1972; *Kunetz*, 1972; *Lee et al.*, 1987; *Levy et al.*, 1988; *Lee et al.*, 1989], seismic forward modeling techniques, as well as the different seismic migration and inverse formulations, have not met a great success in electromagnetism. The method of *Zhdanov and Frenkel* [1983], which proposed the "reverse" propagation of residus from the observers into the medium widely used in seismic migration, has been applied only to homogeneous media with analytical Green functions. This "reverse" propagation requires a process opposite to the diffusion which might be considered as an unstable procedure.

The link between seismic and electromagnetic phenomena is emphasized with the integral transformation between a fictitious wave field and the electromagnetic field. This transformation has been known for a long time [*Filippi and Frisch*, 1969; *Lavrent'ev et al.*, 1980]. An illustration of the forward problem has been given recently by *Lee et al.* [1989]: they first compute the fictitious wave field response to a given conductive medium and transform it into an electromagnetic response, which is a stable procedure. Although they apply finite difference codes, they suggest the use of efficient seismic modeling tools as ray tracing programs for computing the response of a conductive medium.

A more ambitious target is extracting the fictitious wave field from the electromagnetic response. The encountered difficulties are similar to the ones met in the Prony's method or in exponential fittings [*McDonough and Huggins*, 1968]. Construction of an exact inversion operator leads to problems of convergence and stability, the analysis of which is beyond the scope of this paper. How can we transform this ill posed inverse problem into a better posed problem with a tractable solution? From the statistical point of view a stable solution will be the stochastic inverse [*Tarantola*, 1987]. How to obtain this solution is the question we want to address in this paper.

Paper number 91JB00278.
0148-0227/91/91JB-00278$05.00

Once obtained, the fictitious wave field can be processed by seismic migration methods, opening to us the possibility of reconstruction of conductivity inside a given medium.

After a brief description of electromagnetic propagation in conductive media and the integral transformation we set up the inverse problem and describe the difficulties we must solve. This leads us to preprocess our data in order to find our parameterization. Then, the solution is constructed by simulated annealing. With the help of one-dimensional examples, we discuss the efficiency of this technique, the stopping criterion, and the structure of the final solution. Finally, two-dimensional medium geometry is investigated with inversion of synthetic finite element data.

Diffusion of Electromagnetic Response in Conductive Media

If we assume the displacement current is negligible, the electric field $\mathbf{E}$ satisfies the diffusion equation,

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},t) + \mu\sigma(\mathbf{r})\frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t} = \mathbf{S}(\mathbf{r},t) \quad (1)$$

where σ is the electric conductivity of the medium. μ is the magnetic permeability which is taken constant. The source term $\mathbf{S}$ is causal. Following *Lee et al.* [1989], we introduce a fictitious wave field $\mathbf{U}$ and a source term $\mathbf{F}$ related by the following relation:

$$\nabla \times \nabla \times \mathbf{U}(\mathbf{r},q) + \mu\sigma(\mathbf{r})\frac{\partial^2 \mathbf{U}(\mathbf{r},q)}{\partial q^2} = \mathbf{F}(\mathbf{r},q) \quad (2)$$

where q is an independent variable with the dimension of square root of time. The corresponding "seismic" velocity turns out to be $1/\sqrt{\mu\sigma}$. An integral relation between $\mathbf{E}(t)$ and $\mathbf{U}(q)$ (see *Lee et al.* [1989] for references),

$$\mathbf{E}(t) = \frac{1}{2\sqrt{\pi t^3}} \int_0^\infty q e^{-q^2/4t} \mathbf{U}(q)\, dq \quad (3)$$

introduces a kernel corresponding to the diffusive green function of three-dimensional medium. The spatial dependence of $\mathbf{E}$ and $\mathbf{U}$ has been left implicit. This transformation involves only the time t and the variable q and is independent of $\mathbf{r}$. The source terms $\mathbf{S}$ and $\mathbf{F}$ are related by the same expression. One can write the expression (3) in the spectral domain under a simpler form,

$$\mathbf{E}(\omega) = \int_0^\infty e^{-\sqrt{i\omega}\, q}\, \mathbf{U}(q)\, dq \quad (4)$$

already used by *Weidelt* [1972]. We select the following square root of i

$$\sqrt{i} = \frac{1}{\sqrt{2}}(i+1), \quad (5)$$

with an exponentially damped kernel, making the inverse problem difficult to solve. Following *Kunetz* [1972] and *Levy et al.* [1988], we assume that only reflections contribute to $\mathbf{U}$, and we neglect the slow retrodiffusion associated with smooth variations. We look after sharp discontinuities of conductivity, following an approach common in seismic migration problem. Moreover, we assume that the response has been deconvolved of the source function. The excitation function is a plane wave propagating vertically. The associated wave response is an impulsive signal for each reflection,

$$\mathbf{U}(q) = \sum_{n=1}^{N} \mathbf{U}_n \delta(q - q_n), \quad (6)$$

when we neglect response beyond critical distance which is a well-verified assumption except for very steep reflections. Inserting this expression into equation (4), one component of the electric field can be written:

$$E(\omega) = \sum_{n=1}^{N} U_n e^{-i\omega\tau_n}, \quad (7)$$

where τ_n equals q_n^2. In the time domain we obtain a similar expression,

$$E(t) = \frac{1}{2\sqrt{\pi t^3}} \sum_{n=1}^{N} U_n \sqrt{\tau_n} e^{-\tau_n/4t} \quad (8)$$

We want to estimate the time shifts τ_n and the amplitudes U_n, from a given set of M measured data $E(\tau_m)$ or $E(\omega_m)$. Because $E(\omega_m)$ has a simpler form, we select it for our numerical investigations. The number of events is denoted by N.

From our different numerical tests we can argue that the selection of the number N of events is very important for a convergence toward reasonable solutions. Similar conclusions have been previously stated for the Prony's method applied in time series analysis [*Van Blaricum and Mittra*, 1978] where an erroneous number of events leads to unphysical time shifts and amplitudes of reflections. In other words, the number N of reflections, i.e., the number of basic functions, must be also estimated from the data.

In a more theoretical framework [*Hadamard*, 1932] this problem is ill posed because values of some parameters are not continuous functions of the data. Let us assume two models with one reflection of given amplitude defined by:

$$\begin{aligned} E_1(\omega) &= U_1 e^{-\sqrt{i\omega t_1}}, \\ E_2(\omega) &= U_2 e^{-\sqrt{i\omega t_2}}. \end{aligned} \quad (9)$$

The distance between these two models can be defined as the usual modulus

$$d(E_1;E_2) = \mid U_1 - U_2 e^{-(\sqrt{i\omega\tau_2} - \sqrt{i\omega\tau_1})} \mid e^{-\sqrt{\omega\tau_1}/2} \quad (10)$$

For any values of amplitudes U_1 and U_2 we can find arbitrary times τ_1 and τ_2 making this distance as small as we want. From the practical point of view this means that a wide range of models gives identical images in the data space. Fortunately, this difficulty is partially overcome by the observed frequency range and the max-

imum depth that we consider. This often gives a stable solution. In any case, we can increase the stability by the very powerful constraint $| U_n | < 1$. This constraint, which states that reflections have a smaller amplitude than the emitted incident wave, makes the problem well behaving with a global minimum for the misfit function, as we shall see.

Estimation of the Number of Reflections

Selecting the number N of events is a very important step and makes the problem more stable. *Levy et al.* [1988] have proposed on a similar subject to reduce the nonlinear problem of finding amplitudes U_n and times τ_n to a linear problem in amplitudes when the time τ_n of each event is fixed. The solution is regularized by reducing the number of significant events by mean of a minimization of the L_1 norm of the solution vector (see also *Oldenburg* [1990]).

We prefer an alternative approach which we describe now. Let us define one component of the reflected field by:

$$E(\omega_m) = \sum_{n=1}^{N} U_n e^{-\sqrt{i\omega_m \tau_n}} \qquad (m = 1, ..., M), \quad (11)$$

in such a way that the M frequencies are sampled with a regular step in the square root of frequency, which means that

$$\sqrt{\omega_m} = \nu_m \qquad with \qquad \nu_m = m\delta\nu + \nu_0. \quad (12)$$

The reflected field can now be written:

$$E(\nu_m) = E_m = \sum_{n=1}^{N} U_n \, {}_0Z_n Z_n^m, \quad (13)$$

with

$$\begin{aligned} {}_0Z_n &= e^{-\sqrt{i\tau_n}\nu_0}, \\ Z_n &= e^{-\sqrt{i\tau_n}\delta\nu}. \end{aligned} \quad (14)$$

The expression (13) is very similar of the Prony model [e.g., *Kay and Marple*, 1981] widely applied for spectral analysis of transient signals in the terminology of signal processing. N is the order of the model. Among many ways to evaluate the order of Prony's model, we have found the spectral analysis of the data covariance matrix introduced by *Van Blaricum and Mittra* [1978] a very efficient method, which we present in the appendix.

Let us repeat that the decoupling between the number of reflections and the reflection positions provides us with a stable and efficient inversion approach, while inverting simultaneously these quantities is a more unstable procedure.

Simulated Annealing

Once we know how many reflections are expected, we must locate them: this is our nonlinear inverse problem. In order to estimate the misfit function we need the amplitude U_n which is computed by standard techniques [*Lawson and Hanson*, 1974] of linear least squares inverse problems under constraints. Let us underline that amplitudes are not involved in the nonlinear inverse problem as unknowns: they are computed as quantities once the nonlinear parameters, i.e., the reflection positions, are selected.

What makes this inverse problem ill posed? Let us consider the misfit function, hereafter also referred to as the energy function, with L_2 norm defined by

$$\mathcal{S} = 1/2 \sum_{m=1}^{M} (E_{obs}(\omega_m) - E(\omega_m, \tau_1, ..., \tau_n))^2 \quad (15)$$

where we sum over M frequencies sampled in the observed frequency range. For a zero-mean Gaussian white noise with variance σ^2, the energy function may readily be converted to a probability density function

$$\mathcal{P} = \frac{1}{C} e^{-\mathcal{S}/\sigma^2} \quad (16)$$

where C is a normalizing constant. For each assumed position of the reflections we compute the amplitude. The misfit function is deduced from the amplitude misfit.

For a given model with three reflections at positions q = 0.067., 0.139, and 0.204, let us plot the misfit function which is a three-dimensional function in this particular case. We consider slices of this function as shown in Figure 1, and obvious symmetries are related to permutations of the three events. For example, the global minimum is located at six different places. When the third event is at q = 0.204, i.e., between the first and second slice starting from the top of the Figure 1, we have two symmetrical positions coming from permutation of the first two events. Two other positions of the global minimum are also found near the fourth slice from the top of Figure 1 when the third event is at q = 0.139. Finally, we observe two other positions when the third event is at q = 0.067 near the sixth slice from the top. Although the global minimum is well defined, the procedure to reach it by gradient methods can be quite slow, because we have to turn around apparent potential barriers: starting in the sixth slice from the top, we do not converge to the global minimum of this slice if we need for that to go through a situation with two nearby reflections. When an unwanted reflection is near a true one, the inverse procedure balances between these two reflections for interpreting the single reflection and gives a misfit function higher than for a single event, creating a small barrier not seen in the Figure 1. The game is to go to the global minimum using a rather long path preventing the situation of two nearby reflections. Moreover, the main valley of the global minimum presents a rather flat bottom which can be very sensitive to noise perturbation: another reason to avoid using gradient methods. The two sides of the valley have very steep slopes which can be seen, for example, on local variations perpendicular to slices near the fourth slice from the top and, consequently, any discretization of the cost function must take care of apparent local minima coming from the sampling. How to escape from these difficulties?

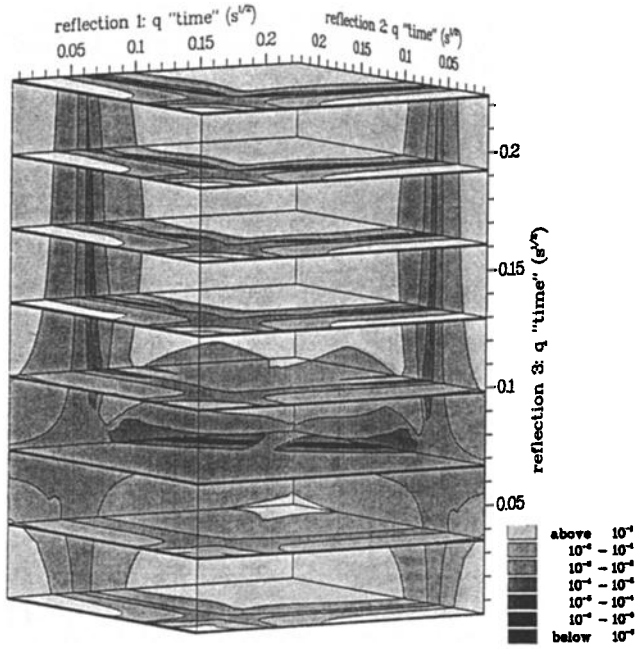

Fig. 1. Misfit function for three reflections. Slices are plotted in grey scale. Variations perpendicular to slices are shown on two background planes. Symmetry of the figures comes from permutations between the three reflections: the global minimum is found above the second slice from the top, above the fourth one, and near the sixth one. For each of these slices, two positions of the global minimum are obtained by permutation of the two other events. The path performed by a gradient method from any arbitrary point in the cube is a rather complex line which must avoid a situation with two reflections nearly at the same position.

The simulated annealing technique, originally proposed by *Kirkpatrick et al.* [1983] and *Cerny* [1985], is the answer we select compared to a pure Monte-Carlo approach. Although this technique requires the solution of the forward problem many times, it has proven its efficiency in many situations (see *van Laarhoven and Aarts* [1987] for a review and *Rothman* [1985, 1986], *Landa et al.* [1989], and *Basu and Frazer* [1990] for geophysical applications). Simulated annealing is a two-loop procedure which is a compromise between local convergent methods and Monte-Carlo methods. The first (inner) loop is the Metropolis algorithm which consists of randomly perturbing the model parameters and accepting the new model with a probability

$$\mathcal{P}_M = \min(1; e^{-\Delta\mathcal{S}/T}), \tag{17}$$

where $\Delta\mathcal{S}$ is the change in the misfit function due to the model perturbation, and T is a parameter called, for historical reasons, the temperature. The probability $\mathcal{P}_M$ is such that when the change $\Delta\mathcal{S}$ is positive (i.e., the new model is worse) the perturbation is not systematically rejected. The second (outer) loop of simulated annealing is the cooling schedule which consists of decreasing the temperature while looping over the Metropolis algorithm. It can be shown that for an infinite number of attempted perturbations the energy distribution of the accepted models obeys the Boltzmann law

$$\mathcal{P}_B = \frac{1}{Z} e^{-\mathcal{S}/T}, \tag{18}$$

where Z is a normalizing partition function. The similarities between the probability distributions $\mathcal{P}$ of (16) and $\mathcal{P}_B$ of (18) allows for a bridge between simulated annealing and inverse problem theory [*Mosegaard and Tarantola*, 1991]. In fact, the temperature T and the noise variance σ^2 play the same role, and decreasing the temperature during the cooling schedule is equivalent to gradually enhancing the influence of the data (i.e., increasing the signal-to-noise ratio) when selecting the models in the Metropolis loop.

In practice, the Metropolis algorithm is not efficient at low temperature where the rejection probability level of most perturbations is high, and it can be replaced by a faster procedure known as the heat bath method [*Creutz*, 1980] and applied by *Rothman* [1986]. For each of the k possible values of the currently considered time shift τ_n (keeping all the other time shift parameters fixed) one computes the relative probability of acceptance by the following expression,

$$\mathcal{P}_A(\tau_n, k) = \frac{e^{-\mathcal{S}_k/T}}{\sum_{k=1}^{K} e^{-\mathcal{S}_k/T}}, \tag{19}$$

for each k between 1 and K. The denominator of (19) may be viewed as a marginal partition function. From this probability distribution one makes a random guess. This guess is always taken and gives the new value for τ_n. We must repeat this procedure for every reflection.

Whether simulated annealing is efficient or not strongly depends on the possibility of the process to jump from one particular acceptable model to another one. When the temperature is high (i.e., when the singal-to-noise ratio is assumed low), almost every particular model is likely to be accepted. However, when the temperature decreases (i.e., when the signal-to-noise ratio is assumed larger and larger), only a few models, belonging to a relatively small subset of the model set, remain with a significant likelihood of acceptance. For such low temperatures, and due to the local nature of the attempted transitions, it may be difficult, in practice, to tunnel from an acceptable model to another one since the number of worse and intermediate models to visit may be large [*Hajek*, 1988]. Obviously, this number changes according to the manner in which we perform our model perturbations. It has been reported previously that the tunneling effect is sometimes easier when the perturbation involves several model parameters simultaneously [*Ettelaie and Moore*, 1987; *Heynderickx and De Raedt*, 1988]. Our numerical attempts showed us that these more sophisticated perturbation schemes have no effects in our problem. Consequently, at each iteration, we have perturbed a single parameter, namely, the position of one reflection.

In summary, our algorithm proceeds in the following steps: selection of the number of events N as explained in the previous section and computation of the acceptance probability (equation (19)) used for the drawing

of a new value of the selected position. When we have considered every reflection, we have performed a single iteration. The simulated annealing repeats this iteration while the temperature is decreased in an adequate way.

Cooling Schedule

Good control over the convergence parameter T is the difficulty of the simulated annealing. By a too slow cooling, we can make the solution very expensive, while a too quick cooling might trap us into an unwanted solution [*Kirkpatrick et al.*, 1983; *Kirkpatrick*, 1984]. In other words, the convergence rate toward a minimum increases when the cooling rate increases, but the probability to be in the global minimum diminishes. Therefore, we must have an adequate slow cooling, and until now, despite theoretical attempts [*Hajek*, 1988], only experiences in different practical situations [*Rothman*, 1985] provide bounds to this cooling.*Nulton and Salamon* [1988] and *Andresen et al.* [1988] propose an adaptative annealing schedule for global optimization problems.

In our case, we have selected the strategy proposed by *Huang et al.* [1986] where the cooling is made at a constant thermodynamic speed λ. We must verify that the averaged energy at iteration $n+1$ is below the averaged energy at the iteration n by λ times the standard deviation, denoted sd( ), of the energy at iteration n. Getting realistic estimates of these statistical quantities implies that the energy distribution at a given iteration must be sampled accurately before changing the temperature. We have the following formula,

$$< \mathcal{S}(n+1) >=< \mathcal{S}(n) > -\lambda\ sd(\mathcal{S}(n)), \qquad (20)$$

where λ is chosen in order to guarantee an appreciable overlap between the energy distributions at iterations n and $n+1$. The greater this overlap, the larger the probability to escape from local minima. The cooling law,

$$T(n+1) = T(n)e^{-\lambda T(n)/sd(\mathcal{S}(n))}, \qquad (21)$$

is deduced [*Huang et al.*, 1986]. The choice of the parameter λ is not crucial and we have taken $\lambda = 0.75$, a rather high value compared to those ($\sim$ 0.1) often chosen by most authors. In practice, the thermal quasi-equilibrium at a given temperature is quickly reached in accordance with the results obtained by *Creutz* [1980], and 10 iterations at constant temperature are sufficient to give good estimates of average energy and standard deviation.

Finally, we must specify the initial temperature and the stopping temperature of the cooling procedure. The initial temperature must be high enough to allow nearly any transitions (i.e., the probability distribution $\mathcal{P}_A$ of (19) is almost uniform). The temperature is taken as the averaged energy over 100 random configurations incremented of the standard deviation :

$$T_{initial} = < \mathcal{S}(n) > + sd(\mathcal{S}(n)). \qquad (22)$$

This choice allows more than 80% of possible transitions, as suggested by *Kirkpatrick* [1984]. The selection of the final temperature is less obvious. A temperature equal to the noise variance,

$$T_{noise} = sd(\text{noise})^2, \qquad (23)$$

might be good since for this particular choice the probability distributions $\mathcal{P}$ (equation (16)) and $\mathcal{P}_B$ (equation (18)) are identical, and the model space is sampled according to the a posteriori distribution $\mathcal{P}$ [*Mosegaard and Tarantola*, 1991]. By decreasing the temperature below this level it can be shown that each parameter converges towards the maximum-likelihood solution [*Mosegaard and Tarantola*, 1991]. This encourages us to continue the cooling until the following final temperature:

$$\frac{T_{noise}}{100} < T_{final} < \frac{T_{noise}}{10}. \qquad (24)$$

Finally, increasing in one step the temperature to T_{noise} will allow the computation of the conditional probability distributions $\mathcal{P}_A$ (equation (19)) which are our final solution and give an insight upon the location error of the reflection.

TABLE 1. Model Description

Conductivity, s/m	Layer thickness, m
0.01	300.
0.10	100.
0.02	100.
0.01	150.
0.10	∞

Simple Example for Testing Simulated Annealing

The selected example is a one-dimensional medium with four reflections, described in Table 1. The positions of primary reflections are converted in travel time τ and variable q in Table 2. The amplitudes of reflections are also given in Table 2. We compute electromagnetic responses using (3) at 100 frequencies between 1 Hz and 1.5 kHz. We add a Gaussian white noise with zero mean to this synthetic data. We construct three data sets with signal-to-noise ratio of 30 dB, 20 dB and 10 dB (Figure 2). The spectral analysis of data covariance matrices using the Akaike's information criterion -AIC- (see Table 3) shows us that three reflections can be located for the first two data sets, while the presence of important noise in the last one allows only two reflections to be located. The statistical criterion is constructed for a maximum of 15 expected reflections ($= N'$), a number significantly higher than the a priori selected number of reflections (N), always lower than

TABLE 2. Positions and Amplitudes of Events

Events	Amplitude U	τ, ms	q, $\sqrt{s}$
1	–0.52	4.524	0.067
2	0.28	19.088	0.139
3	0.11	28.860	0.170
4	–0.31	41.420	0.204

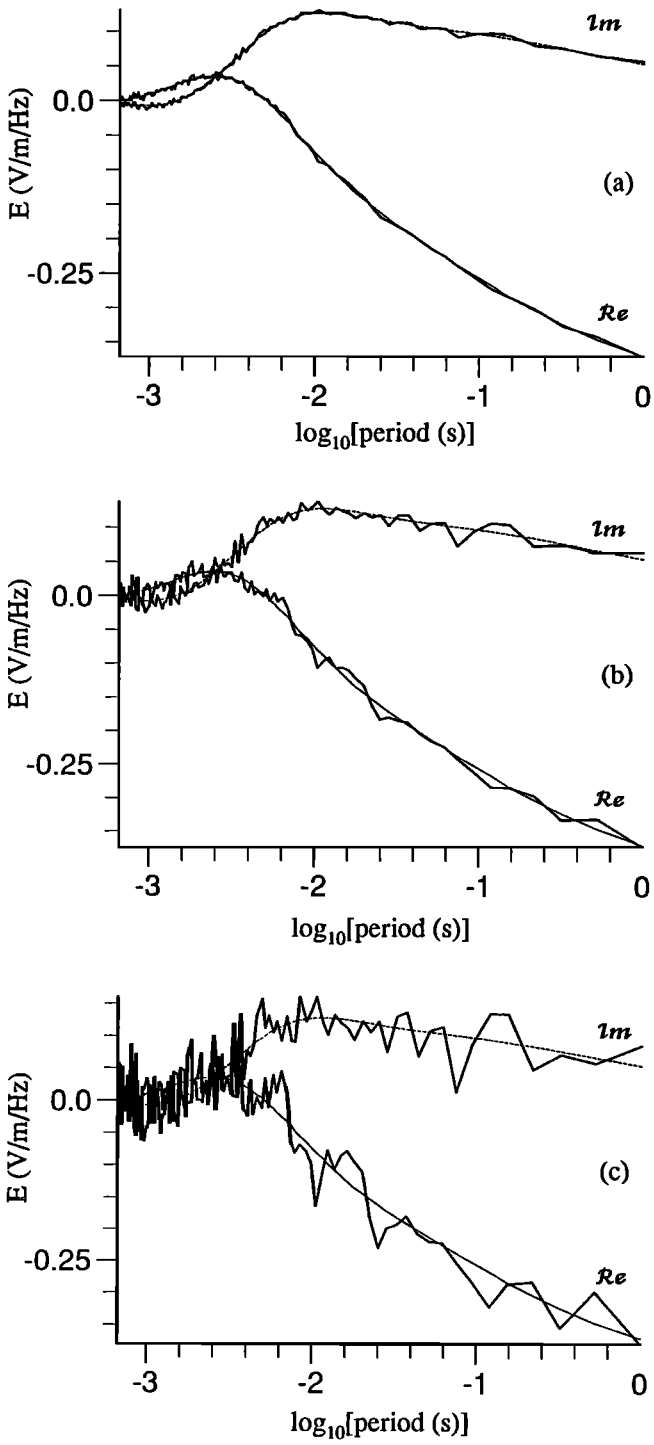

Fig. 2. Synthetic data obtained by convolution. Three signal-to-noise ratios are defined: (*a*) a ratio of 30 dB (variance $\sigma^2 = 10^{-5}$), (*b*) a ratio of 20 dB (variance $\sigma^2 = 10^{-4}$) and (c) a ratio of 10 dB (variance $\sigma^2 = 10^{-3}$).

six events. This is necessary in order to have a significant number of terms in the sum and product of (A9) (see the appendix).

For the high signal-to-noise ratio of 30 dB, the resolvable three reflections given by the criterion AIC are well located (Figure 3) when T_{final} equals $T_{noise}/10$.

TABLE 3. AIC Value for an a Priori Number of Events

Events	Signal-to-Noise ratio, dB		
	30	20	10
1	1388	345	289
2	331	321	**277**
3	**318**	**315**	299
4	334	331	322
5	361	359	348
6	382	380	366

Bold numbers give the AIC selection for the number of events.

The amplitudes are also well defined (Table 4). Let us note that the two intermediate events have the same sign in amplitude and are not resolved by our procedure. The estimated amplitude of the second reflection is nearly the sum of these two intermediate real reflections. Once the positions are given, we compute the conditional probability at T_{noise} (Figure 3). This probability would lead to a possible analysis of the final resolution which is not performed in this study. Intuitively, one can see in Figure 3 the decrease in resolution with depth.

It is interesting to follow the evolution of the event locations when we proceed to the cooling (Figure 4). In the same figure we give the temperature evolution, the decrease of the averaged energy $< \mathcal{S} >$. One can see the hierarchy where the most superficial reflection is first located, the second deeper one, and so on. This evolution is the one we obtain in all the numerical simulations we undertake. The energy variation are best seen with the computation of the derivative $\Delta < \mathcal{S} > / \Delta T$, which is analogous to a specific heat. The decrease of the energy is associated with a better organized system, which is often related, if not always, to the interface position.

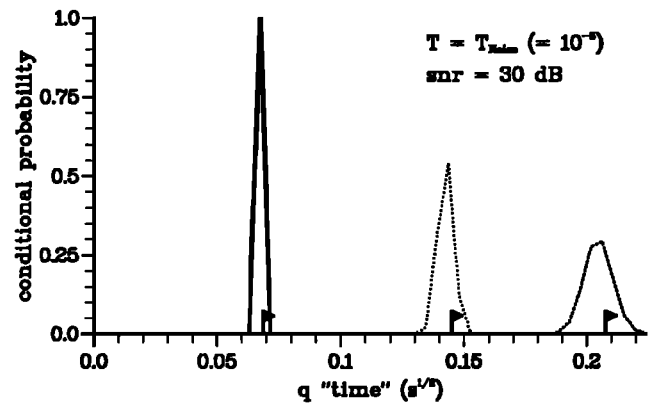

Fig. 3. For a signal-to-noise ratio of 30 dB, conditional probabilities associated to each of the inverted events. Flags indicate final positions of the reflections obtained at the end of cooling schedule ($T = T_{final}$ from equation (24)). For each reflection, conditional probabilities have been computed by varying the position of the considered event (see equation (19)) while keeping the remaining reflections at their final positions.

TABLE 4. Amplitude Estimation of Events After Simulated Annealing

Events	True Amplitude	Figure 3	Figure 5	Figure 6	Figure 7	Figure 8
1	−0.520	−0.520	−0.520	−0.525	0.001	−0.828
2	0.280	0.353	−0.009	0.376	−0.525	0.451
3	0.110	−0.272	0.381	−0.288	0.463	–
4	−0.310	–	−0.291	–	−0.375	–

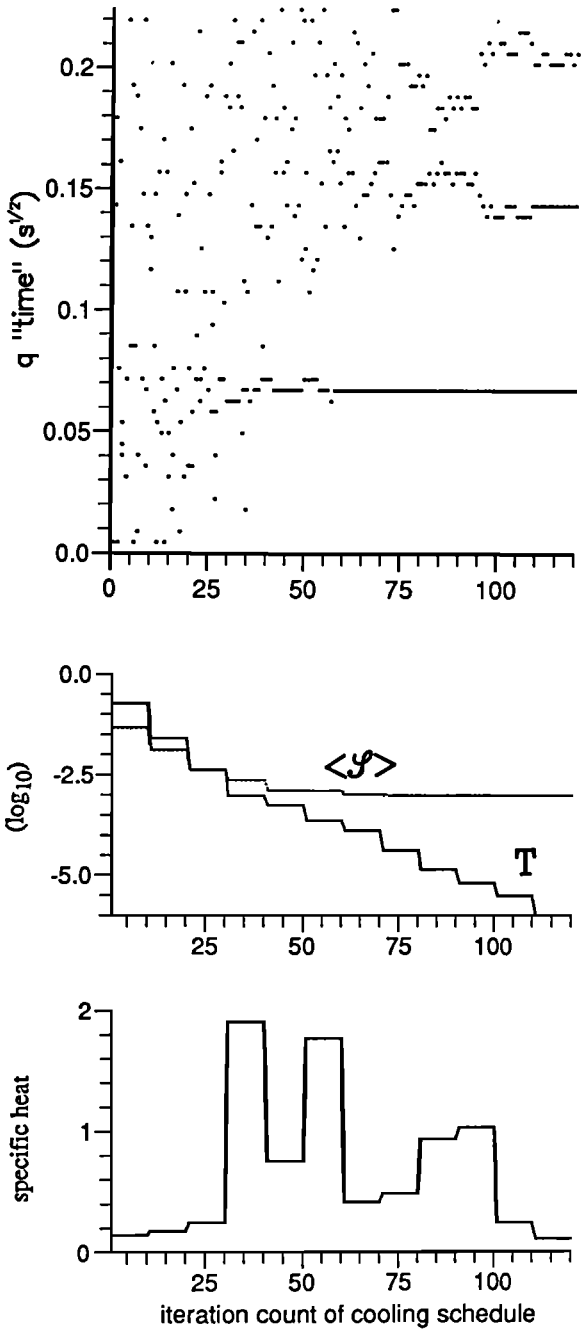

Fig. 4. Evolution of reflection positions on the top panel during the cooling. Intermediate panel shows the prescribed temperature evolution and the decreasing energy. Variations of the energy are better seen with the specific heat coefficient on the bottom panel.

On the bottom of Figure 4, the specific heat presents three bumps related to the temperature interval where an interface starts to freeze at a given position.

When four reflections are a priori sought, the unresolved event, which is the second one with a small amplitude (Table 4), perturbs the positions of the three other reflections and decreases the global resolution (Figure 5). This is an a posteriori proof of the importance of detecting the number N of events and gives a justification for the AIC we have used.

Of course, when the noise increases, events are much more difficult to locate. Figure 6 presents three located pulses at the final temperature $T_{noise}/10$ with a strong degradation of resolution: only the first event is perfectly resolved. Introducing an extra event with respect to the AIC still gives a nonlocated event which perturbs the two deepest events both in locations (Figure 7) and in amplitudes (Table 3). We emphasise that this event has a very small amplitude which can be used to reject it.

For a noisy environment as the final example, the *AIC* allows two reflections. One can see that the first one is welllocated, but the second one is located in a wrong position near the first one at the final temperature (Figure 8). The whole strategy fails for this signal-to-noise ratio, and the second reflection interprets only noise in the data (Figure 2*c*).

TWO-DIMENSIONAL REALISTIC EXAMPLE

In this last section we would like to test the performance of simulated annealing on synthetic data obtained by an entirely different forward modeling than the integral equation (3). We have used a finite element

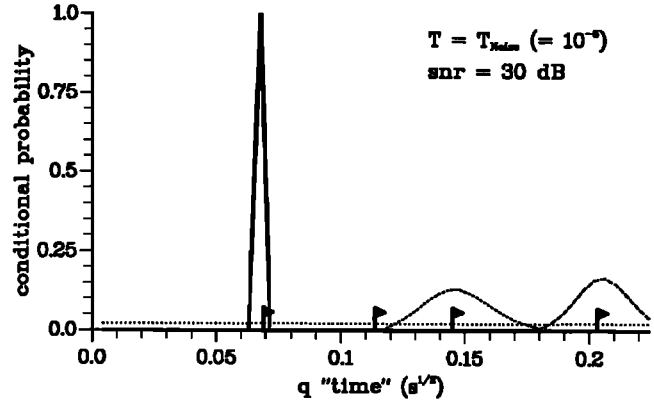

Fig. 5. Same as Figure 3 for a priori four reflections. Please note the unlocalized event.

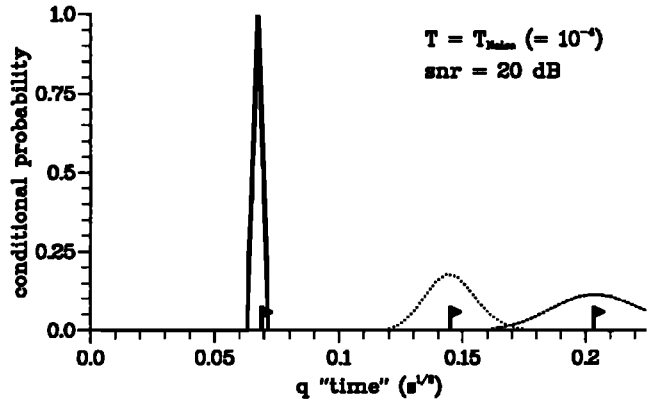

Fig. 6. Same as for Figure 3 but with a signal-to-noise ratio of 20 dB. Three reflections are deduced from the criterion AIC.

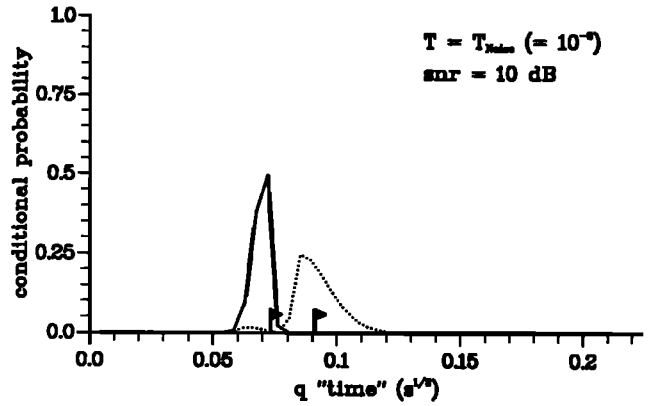

Fig. 8. Same as for Figure 3 but for a priori two events and a signal-to-noise ratio of 10 dB. One reflection is correctly located, while the second is at a wrong position and fits some noise in the signal.

program [*Wannamaker et al.*, 1985] for a simple realistic medium. This model is designed such that the deepest and third interface is flat with an important constrast of the resistivity. Above this interface a more complex pattern of interfaces is considered first for perturbing the diffusion process and for checking the performance of the simulated annealing. Moreover, resistivity contrasts are smaller than for the deepest interface (Figure 9). We neglect displacement current as usually done in the usual magnetotelluric approximation, and we consider transverse magnetic mode for our experiment.

We use 63 traces with only 50 frequencies between 1 Hz and 1.5 kHz with the regular sampling of the square root of frequency (equation (12)). An extra Gaussian white noise is added with a signal-to-noise ratio of 30 dB.

During the cooling for each trace, the parameter λ is equal to 0.75, while 10 iterations are performed at each temperature before the selection of the next position. The final temperature is taken as $T_{noise}/10$. The criterion AIC turns out to select between three and six events, although only two reflections are correctly located: the final conditional probabilities of these extra events are rather flat.

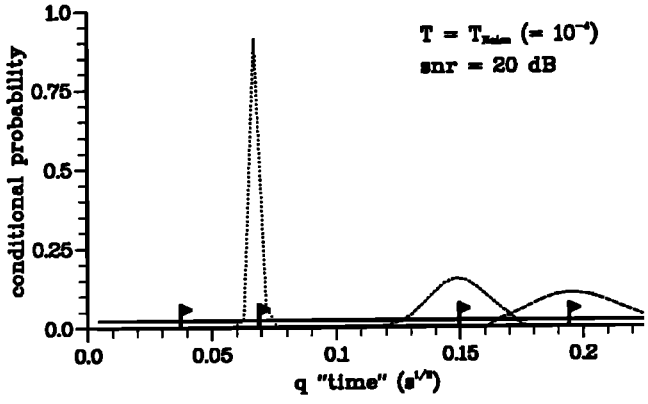

Fig. 7. Same as for Figure 3 but for a priori four events and a signal-to-noise ratio of 20 dB. Please note the unlocalized event and the unwanted perturbation of the latest reflection.

The conditional probability is shown for each trace in an equivalent seismic section in Figure 9. We emphasize that the electromagnetic source is a vertical plane wave exciting the medium at the free surface. In other words, the pseudo-seismic section is the one expected after a CMP stack before migration.

The simulated annealing has the capacity to detect two interfaces: The top interface is the best localized, while the second is slightly blurred out. This second interface, which is flat in the real medium, has a slope in good correlation with the resistivity above it. The complexity on the right of the model is entirely unresolved and does not even degrade the solution: interfaces are too near each other with a too small jump of resistivity.

Sometimes, reflections are mislocated with a spread conditional probability: a posteriori analysis might require the restarting of the cooling schedule, and an improved picture would be obtained. We have not undertaken this strategy and have preferred to present straight results from the inversion.

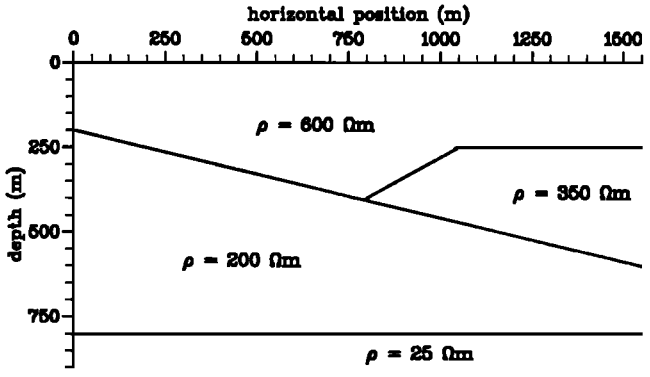

Fig. 9. The two-dimensional model on the top with a deep flat interface. From TM signals at the free surface, we deduce a pseudoseismic section which is equivalent to a CMP stack before migration. More accurate readability of the event position is obtained by plotting the conditional probability at the temperature T_{noise}.

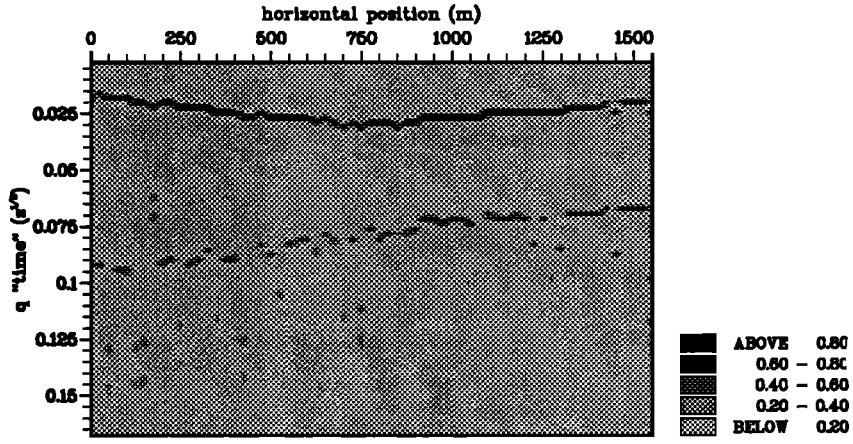

Fig. 9. (continued)

The conversion from pseudotime to depth will require the use of migration techniques used in seismic data analysis. The discussion of this transformation is beyond the scope of this paper. Poststack migration might be applied to this pseudo-seismic section and, hopefully, the deepest interface will become flat.

Conclusion

We have shown that diffusive electromagnetic signals can be converted to impulsive pseudo-wave signals by a procedure similar to deconvolution. This procedure is efficiently performed by simulated annealing. We have studied the different steps in order to achieve this transformation and have defined the parameters controlling the cooling. Except for the very noisy example, we succeed in locating reflections, as well as in estimating their amplitudes.

This success enables us to propose the different migrations of seismic exploration in order to construct the contrast of conductivity inside the medium. This will be the purpose of a future work.

Each trace is inverted independently from the neighboring ones. A promising strategy for future work will include the coherence between traces to accelerate the search for the global minimum for each trace. This technique will improve the pseudo-seismic section we have obtained in the previous paragraph.

Appendix

The spectral analysis of the data covariance matrix introduced by *Van Blaricum and Mittra* [1978] turns out to be a very efficient method for determining the number of reflections.

Let us define the polynomial $P(Z)$ of total degree N with roots Z_n,

$$P(Z) = \prod_{n=1}^{N}(Z - Z_n) = \sum_{\ell=0}^{N} a_\ell Z^{N-\ell}, \qquad (A1)$$

with a_0 equal to 1. Multiplying (13) by a_ℓ, we obtain the following expression:

$$a_\ell E_{j-\ell} = \sum_{n=1}^{N} a_\ell U_{n\ 0} Z_n Z_n^{j-\ell}, \qquad (A2)$$

with $\ell = 0, \cdots, N$ and $j = N+1, \cdots, M$. By summing the N terms on the LHS, we obtain the expression,

$$\sum_{\ell=0}^{N} a_\ell E_{j-\ell} = \sum_{n=1}^{N} U_{n\ 0} Z_n \sum_{\ell=0}^{N} a_\ell Z_n^{j-\ell}, \qquad (A3)$$

with $(j = N+1, \cdots, M)$. The polynomial on the RHS is evaluated at one of its roots Z_n and is consequently equal to zero. Thus the linear system

$$\sum_{\ell=0}^{N} a_\ell E_{j-\ell} = 0 \qquad (A4)$$

with $a_0 = 1$ and $j = (N+1, \cdots, M)$, holds and can be written in the compact form: $\mathcal{E}\mathbf{a} = 0$, where we have introduced the matrix $\mathcal{E}$,

$$\mathcal{E} = \begin{bmatrix} E_1 & \cdots & E_{N+1} \\ \vdots & \ddots & \vdots \\ E_{M-N} & \cdots & E_M \end{bmatrix}, \qquad (A5)$$

and the vector $\mathbf{a}^t = (a_N, \cdots, a_0)$. As usual, we reduce to a square system by multiplying (A4) with the conjugated transposed matrix $\mathcal{E}^h$ of $\mathcal{E}$. If we divide also (A4) by the number of unknown parameters $N+1$, we obtain the data covariance matrix $\mathcal{R}$ and the equivalent system,

$$\mathcal{R}\mathbf{a} = \mathbf{0}. \qquad (A6)$$

This expression shows us that $\mathbf{a}$ is an eigenvector of the matrix $\mathcal{R}$ with an eigenvalue equal to zero. The rank of matrix $\mathcal{R}$ is lower or equal to the number of events N. Of course, in presence of noisy data, eigenvalues are not identically zeroes: we must select a threshold under which we decide with a given criterion that the eigenvalue is almost zero. For white noise with variance σ^2, the covariance matrix has the following form with $\mathcal{I}$ as the identity matrix,

$$\mathcal{R}' = \mathcal{R} + \sigma^2 \mathcal{I}, \qquad (A7)$$

from which one can obtain:

$$\mathcal{R}'\mathbf{a} = \sigma^2\mathbf{a} \qquad (A8)$$

which states that the threshold can be taken as σ^2 [*Van Blaricum and Mittra*, 1978; *Kay and Marple*, 1981]. As statistical test of decision, we use the Akaike's test selected by *Wax and Kailath* [1985] for a very similar problem. We compute a data covariance matrix for a number of reflection N' higher than the expected number N. The true number of reflections is the one which makes the Akaike's information criterion -AIC- minimum:

$$\mathrm{AIC}(N) = -2\mathrm{Log}\left[\frac{\prod_{i=N+1}^{N'} \lambda_i^{\frac{1}{N'-N}}}{\frac{1}{N'-N}\sum_{i=N+1}^{N'} \lambda_i}\right]^{(N'-N)(M-N')} + 2N(2N'-N), \qquad (A9)$$

where the λ_i are the eigenvalues of $\mathcal{R}'$ arranged in decreasing order. We found this criterion very successful in obtaining the number of reflections in even relatively noisy signals. For a more detailed description of this criterion we refer the reader to *Wax and Kailath* [1985].

Acknowledgments. Many thanks to D. W. Oldenburg for making available his preprints and for helpful comments. This paper benefited from stimulating discussions with C. Hubans, V. Jouanne, M. Menvielle, P. Podvin, M. Roussignol, J. T. Smith, D. Sornette, and P. Tarits. Very constructive reviews were made by K. Mosegaard and A. Tarantola. This research was supported by the sponsors of CGGM de l'Ecole des Mines (CFP, CGG, IFP, SNEAP).

REFERENCES

Andresen, B., K.H. Hoffmann, K. Mosegaard, J.D. Nulton, J.M. Pedersen, and P. Salamon, On lumped models for thermodynamic properties of simulated annealing problems, *J. Phys. Paris, 49*, 1485-1492, 1988.

Barnett, C. T., Simple inversion of time-domain electromagnetic data, *Geophysics, 49*, 925-933, 1984.

Basu, A., and L.N. Frazer, Rapid determination of the critical temperature in simulated annealing inversion, *Science, 249*, 1409-1412, 1990.

Cerny, V., A thermodynamical approach to the travelling salesman problem, *J. Optimization Theory Appl., 45*, 41-51, 1985.

Creutz, M., Monte Carlo study of quantized SU(2) gauge theory, *Phys. Rev. D, 21*, 2308-2315, 1980.

Ettelaie, R., and M. A. Moore, Zero-temperature scaling and simulated annealing, *J. Phys. Paris, 48*, 1255-1263, 1987.

Filippi, P., and U. Frisch, Relation entre l'équation de la chaleur et l'équation des ondes de Helmholtz, *Comptes Rendus Acad. Sci., 268-A*, 804-807, 1969.

Hadamard, J., *Le problème de Cauchy et les équations aux dérivées partielles hyperboliques*, Hermann, Paris, 1932.

Hajek, B., Cooling schedules for optimal annealing, *Math. Oper. Res., 13*, 311-329, 1988.

Heynderickx, I., and H. De Raedt, Calculation of the director configuration of nematic liquid crystals by the simulated anneal method, *Phys. Rev. A, 37*, 1725-1730, 1988.

Huang, M.D., F. Romeo, and A.L. Sangiovanni-Vincentelli, An efficient general cooling schedule for simulated annealing, *Proc. IEEE Int. conf. Computer-Aided Design*, Santa Clara, 381-384, 1986.

Kay, S. M., and S. L. Marple, Jr., Spectrum analysis: a modern perspective, *Proc. IEEE, 69*, 1380-1419, 1981. Kirkpatrick, S., Optimization by simulated annealing: Quantitative studies, *J. Statist. Phys., 34*, 975-986, 1984.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, *Science, 220*, 671-680, 1983.

Kunetz, G., Processing and interpretation of magnetotelluric soundings, *Geophysics, 37*, 1005-1021, 1972.

Landa, E., W. Beydoun, and A. Tarantola, Velocity model estimation from prestack waveforms: Optimization by simulated annealing, *Geophysics, 54*, 984-990, 1989.

Lavrent'ev, M. M., V. G. Romanov, and S. P. Shishatskii, *Ill-posed Problems of Mathematical Physics and Analysis* (in Aussian), Nauka, Moscow, 1980.

Lawson, C. L., and R. L. Hanson, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, N.J., 1974.

Lee, K. H., G. Liu, and H. F. Morrison, A new approach to modeling the electromagnetic response of conductive media, *Geophysics, 154*, 1180-1192, 1989.

Lee, S., G. A. McMechan, and C. L. Aiken, Phase-field imaging: The electromagnetic equivalent of seismic migration, *Geophysics, 52*, 1678-693, 1987.

Levy, S., D. W. Oldenburg, and J. Wang, Subsurface imaging using magnetotelluric data, *Geophysics, 53*, 104-117, 1988.

McDonough, R.N., and W.H. Huggins, Best least-squares representation of signals by exponentials, *IEEE Trans. Autom. Control, AC-13*, 408-412, 1968.

Mosegaard, K. and A. Tarantola, Monte-Carlo analysis of inverse problems, *J. Geophys. Res.*, in press, 1991.

Nulton, J.D., and P. Salamon, Statistical mechanics of combinatorial optimization, *Phys. Rev. A, 37*, 1351-1356, 1988.

Oldenburg, D.W., Inversion of electromagnetic data : an overview of new techniques, *Geophys. Surv.*, in press, 1990.

Oristaglio, M. L., and G. W. Hohmann, Diffusion of electromagnetic fields into a two-dimensional earth: a finite difference approach, *Geophysics, 49*, 870-894, 1984.

Rothman, D. H., Nonlinear inversion, statistical mechanics, and residual statics estimation, *Geophysics, 50*, 2784-2796, 1985.

Rothman, D. H., Automatic estimation of large residual statics corrections, *Geophysics, 51*, 332-346, 1986.

SanFilipo, W. A., and G. W. Hohmann, Integral equation solution for the transient electromagnetic response of a three-dimensional body in a conductive half-space, *Geophysics, 50*, 798-809, 1985.

Smith, J. T., Rapid inversion of multi-dimensional magnetotelluric data, Ph.D thesis, Univ. of Wash., Seattle, 1988.

Tarantola, A., *Inverse Problem Theory*, Elsevier Science, New York, 1987.

Tarits, P., *Contribution des sondages électromagnétiques profonds á l'étude du manteau supérieur terrestre*, Thèse, Univ. Paris 7, 1989.

Tripp, A. C., G. W. Hohmann, and C. M. Swift, Two-dimensional resistivity inversion, *Geophysics, 49*, 1708-1717, 1984.

Van Blaricum, M. L., and R. Mittra, Problems and solutions associated with Prony's method for processing transient data, *IEEE Trans. Antennas Propag., AP-26*, 174-182, 1978.

van Laarhoven, P. J. and E. H. Aarts, *Simulated Annealing: Theory and Applications*, D. Reidel, Dordrecht, Holland, 1987.

Wannamaker, P.E., G. W. Hohmann, and W. A. SanFilipo, Electromagnetic modeling of three-dimensional bodies in layered earths using integral equations, *Geophysics, 49*, 60-74, 1984.

Wannamaker, P.E., J.A. Stodt, and L. Rijo, PW2D finite element program for solution of magnetotelluric responses of two-dimensional Earth resistivity structure, Earth Sci. Lab., Univ. of Utah Res. Inst., Salt Lake City, 1985.

Wax, M., and T. Kailath, Detection of signals by information theoretic criteria, *IEEE Trans. Acoust. Speech Signal Process., ASSP-33*, 387-392, 1985.

Weidelt, P., The inverse problem of geomagnetic induction, *J. Geophys., 38*, 257-289, 1972.

Weidelt, P., Inversion of two-dimensional conductivity structures, *Phys. Earth Planet. Inter., 10*, 282-291, 1975.

Zhdanov, M. S., and M.A. Frenkel, The solution of the inverse problems on the basis of analytical continuation of the transient electromagnetic field in reverse time, *J. Geomagn. Geoelectr., 35*, 747-765, 1983.

D. Gibert, Ecole des Mines, CGGM, 35 rue St Honoré, 77305 Fontainebleau cedex, France.

J. Virieux, Institut de Géodynamique, CNRS, Av. Albert Einstein, 06560 Valbonne, France.

(Received May 22, 1990;
(revised January 18, 1991;
(accepted January 22, 1991.)