



Inversion of generalized relaxation time distributions with optimized damping parameter



Nicolas Florsch ^{a,b,e,*}, André Revil ^{c,d}, Christian Camerlynck ^{a,e}

^a Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris, France

^b UMI 209 "UMMISCO", Institut de Recherche pour le Développement, France

^c Colorado School of Mines, Dept of Geophysics, Golden, CO, USA

^d ISTERre, CNRS, UMR CNRS 5275, Université de Savoie, 73376 cedex, Le Bourget du Lac, France

^e UMR-CNRS 7619 « METIS », Paris, France

ARTICLE INFO

Article history:

Received 2 April 2014

Accepted 21 July 2014

Available online 02 August 2014

Keywords:

Induced polarization

Inversion

Relaxation Time Distribution

L-curve

SIP

Impedance spectroscopy

ABSTRACT

Retrieving the Relaxation Time Distribution (RTD), the Grains Size Distribution (GSD) or the Pore Size Distribution (PSD) from low-frequency impedance spectra is a major goal in geophysics. The "Generalized RTD" generalizes parametric models like Cole–Cole and many others, but remains tricky to invert since this inverse problem is ill-posed. We propose to use generalized relaxation basis function (for instance by decomposing the spectra on basis of generalized Cole–Cole relaxation elements instead of the classical Debye basis) and to use the L-curve approach to optimize the damping parameter required to get smooth and realistic inverse solutions. We apply our algorithm to three examples, one synthetic and two real data sets, and the program includes the possibility of converting the RTD into GSD or PSD by choosing the value of the constant connecting the relaxation time to the characteristic polarization size of interest. A high frequencies (typically above 1 kHz), a dielectric term is taken into account in the model. The code is provided as an open Matlab source as a supplementary file associated with this paper.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In physics and electrochemistry, it is usual to analyze polarization phenomena by representing transfer functions in the frequency domain (e.g. Barsoukov and Macdonald, 2005). More recently in geophysics, spectral impedance meters have been used to explore the ground response in both the time and frequency domains. For frequency-domain induced polarization, the frequency range is typically $[10^{-2} \pm 1, 10^3 \pm 2]$ Hz. The data are generally log-equally sampled and provided in terms of resistivity or conductivity, with amplitude and phase or real part and imaginary parts (i.e., in-phase and quadrature components). These representations correspond to the Bode and Nyquist diagrams, respectively.

Beside the traditional few parameters used in induced polarization (like the chargeability in time-domain induced polarization or the "frequency effect", Telford and Sheriff, 1990), more sophisticated models involve the determination of a set of model parameters given a particular analytical function such as the Cole–Cole, Warburg, Cole–Davidson, or Havriliak–Negami distributions (see many other models in Dias, 2000; Hilfer, 2002). A number of authors have addressed this kind of inversion, for instance Ghorbani et al. (2007), Chen et al.

(2008), Miranda and Lopez-Rivera (2008) for the function. Other recent works do exist in the time domain, like the one by Tarasov and Titov (2007), where the authors address the inversion of the model parameters entering the Cole–Cole model.

A recent perspective for interpreting spectral induced polarization data consists in determining the Relaxation Time Distribution (RTD), which can be derived from the complex spectra of impedance or admittance. For instance Nordsiek and Weller (2008) used a Debye decomposition of the spectra to determine a distribution of relaxation times. In this context, it is generally assumed that the observed spectra result from a continuous linear superposition of individual elementary relaxation responses to the external solicitation. Then the question of retrieving the "Time Relaxation Distribution" makes sense and the RTD can be related in turn to some petrophysical properties of interest as for instance the pore or grain size distributions (see Schwarz, 1962; Pelton et al., 1978; Tong et al., 2006; Revil and Florsch, 2010; Revil et al., 2012, 2014). We will not discuss this point in detail in the present paper. Instead, we will focus on the development of mathematical tools to retrieve the RTD from the spectral induced polarization data. In addition, we point out that this question of inverting the RTD from time-domain or frequency domain spectra concerns various disciplines in Earth sciences and physics including rock mechanics (e.g., Hilfer, 2002 on glasses or Revil et al., 2006 for the pressure solution of silica sands).

* Corresponding author at: Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris France.

E-mail address: nicolas.florsch@upmc.fr (N. Florsch).

2. The question of resolution

The resolution of the solution of the inverse problem in spectral induced polarization is often considered to be poor. This may explain the reason why researchers often consider that only a set of few model parameters of a given analytical model can be inverted (frequently the Cole–Cole model parameters). However, depending on his final objective, a geophysicist may probably find useful to actually have the choice in the possible offered parameterizations of the inverse problem including trying to invert the full RTD.

The low resolution in the inversion of the RTD is inherent to the physics of the problem. In the case of a pure Debye relaxation (which include a single time constant τ_0 and the angular frequency ω), the kernel is written as:

$$K(\omega) = \frac{1}{1 + i\omega\tau_0}. \quad (1)$$

The amplitude corresponds to a bell-shaped function given by:

$$|K(\omega)| = \frac{1}{|1 + i\omega\tau_0|} = \sqrt{\frac{1}{1 + \omega^2\tau_0^2}}. \quad (2)$$

Then, $K(\omega)$ is a “Green function” or “impulse response” (actually in the frequency domain) that must be convolved with a time relaxation distribution as soon as this more elementary relaxation is admitted to follow Debye’s model.

Resolving the RTD is not really different than resolving other geophysical anomalies (like in gravimetry, for instance) and the fundamental question regarding the resolution can be formulated as follows: how different must be two constants $\tau_1 = \tau_0/\chi$ and $\tau_2 = \tau_0 \cdot \chi$, with $\chi > 1$, in such a way that it becomes possible to resolve (to separate) these two constants from a mixed spectrum like:

$$F_\chi(\omega) = \frac{1}{1 + i\omega\tau_0/\chi} + \frac{1}{1 + i\omega\tau_0 \cdot \chi}. \quad (3)$$

When the peaks of the independent spectra are seriously overlapping, only one peak seems visible. At this resolution limit, both first and second derivatives vanish in the middle of the unique peak (at axis of symmetry). However, in the general case here, and due to the complex form of the response which makes the problem very different from other classical power resolution problems like in gravity (for instance), we show below that only the imaginary part owns two clearly distinguishable peaks. Fig. 1 shows the spectrum in terms of amplitude, phase, real and imaginary parts, in three cases: (1) under the resolution ($\chi = 2$, indicative value), (2) at the exact resolution ($\chi \cong 2.41$) (see below), and (3) well resolved $\chi = \sqrt{10} \cong 3.16$ (indicative value).

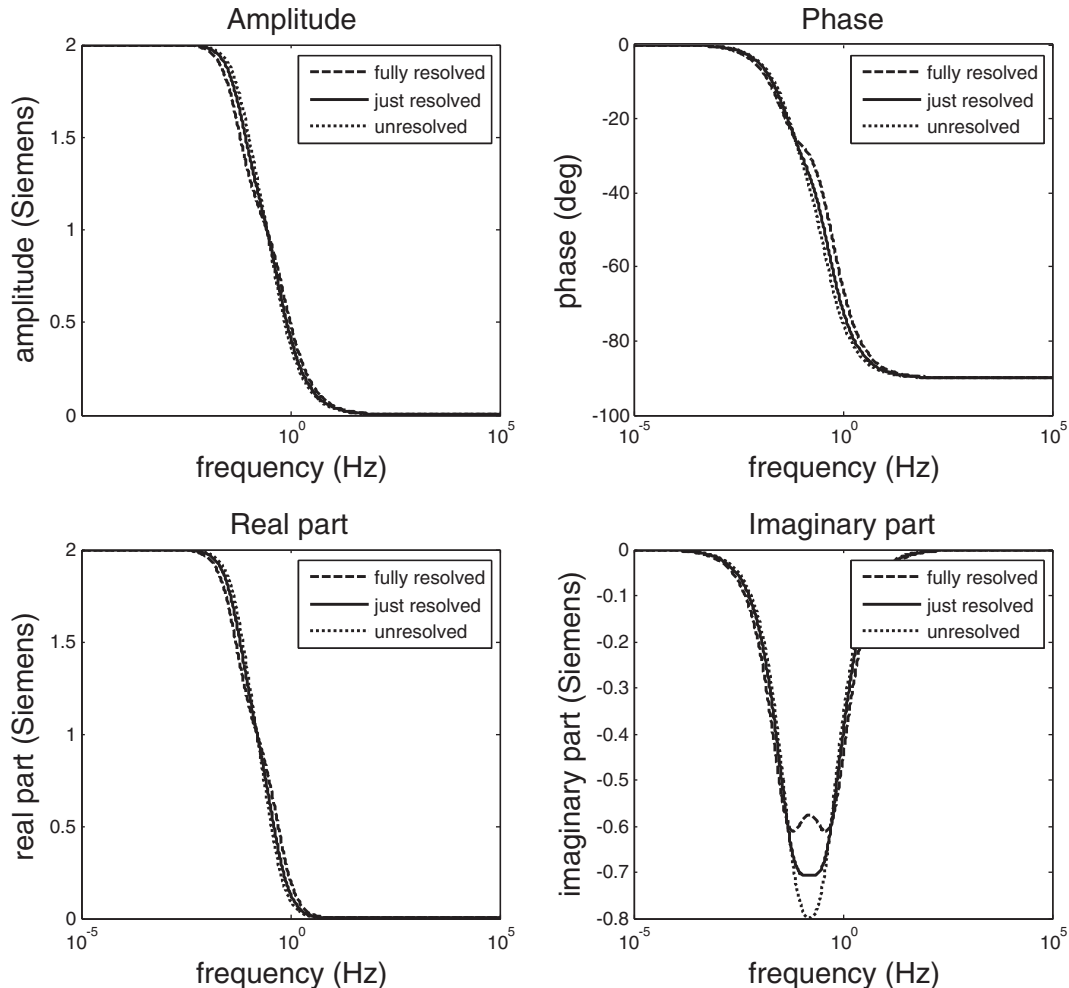


Fig. 1. Only the imaginary part leads to a clear peaks separation, here with $\chi = 2$; $1 + \sqrt{2}$; $\sqrt{10}$ (the last case is one decade).

To look for the exact resolution, one solves:

$$\begin{cases} \frac{dI_{\chi}(\omega)}{d\omega} = 0 \\ \frac{d^2 I_{\chi}(\omega)}{d\omega^2} = 0 \end{cases}, \quad (4)$$

where $I_{\chi}(\omega)$ denotes the imaginary part of $F_{\chi}(\omega)$. After some calculations, one finds $\chi = 1 + \sqrt{2} \approx 2.41$. The two peaks are separated by a factor in the frequency domain equal to the square of χ , which is the ratio:

$$r = \frac{\chi}{1/\chi} = (1 + \sqrt{2})^2 = 3 + 2\sqrt{2} \approx 5.83. \quad (5)$$

The decimal logarithm difference, more important since it is the quantity useful on distribution plots, is given by $\log_{10}(r) = \log_{10}(3 + 2\sqrt{2}) \approx 0.766$. In practice, taking into account some noise and two different amplitudes, it makes sense to consider that the resolution is more or less one decade, that is $r = 10$ (and $\chi = \sqrt{10}$) and a log difference of 1. Therefore, it appears that when attempting to retrieve a continuous spectrum, it is not worth using a sampling higher than four of five points per decade.

The next and more important point concerns the resolution limitations in induced polarization, and the fact that inverting induced polarization spectra belongs to the class of ill-posed problems in the sense of Tikhonov.

3. Materials and method: inverting induced polarization data

Our main objective is to go a step further in inverting and resolving induced polarization spectra in terms of retrieving the RDT and to provide an open source code to the community. This code includes two novelties. The first is relative to the possibility of inverting RTDs based on an inversion of the response using Green functions generalizing the Debye decomposition and convolution form. In other words, our approach offers the possibility to project the spectrum on a generalized Cole–Cole, a Warburg response, or actually any kernel response with some adaptation. The second point is relative to the use of the L-curve, which permits a fair control in the management of the damping factor, which is required to invert an ill-posed problem. The use of the L-curve allows having an optimum trade-off between accuracy and resolution. There are some alternative schemes like the popular Akaike's criteria or cross-validation techniques. However, the use of the L-curve is elegant and powerful and we recently used it successfully to invert the IP data in the Fourier domain (Florsch et al., 2012).

A potential difficulty appearing in inverting SIP data is related to the fact that the distribution of relaxation times needs to be positive-definite. As far as a linear method is used, it occurs that the solution involves some negative contributions, because the positivity is not included as an a priori criterion in the formal mathematics. However, imposing positivity can be done by introducing non-differentiable objective functions or by using a proper bijective and monotone change of variable, like the one given by:

$$x' = 10^x \Leftrightarrow x = \log_{10}(x'); \quad x \in \mathbb{R}; x' \in \mathbb{R}_+. \quad (6)$$

The constraint of positivity makes the problem non-linear. Note that the linear method may yield solutions involving high oscillations (where positive and negative parts compensates), despite the damping, and may lead to unrealistic solutions despite an excellent fit of the data. A possibility to invert a non-linear problem is the Generalized Least-Square approach (abbreviated as TVGLS hereinafter) proposed by Tarantola and Valette (1982). It permits to deal with non-linear objective functions, which results from imposing positivity. A very suitable

change of variables, providing scale-invariance properties, is given by Eq. (6) above. The TVGLS is a form of the Newton–Raphson algorithm, and, as such, it requires an initial guess. The choice of the initial guess is one of the factors determining the success of the inversion. If this initial guess is too far from the (unknown) solution, then instability may occur in the search of the solution depending on the topology of the objective function (e.g., Fernandez-Martinez, 2012). Indeed, far from the solution, the objective function can be flat or random-like, making the steepest gradient approach hopeless. Making a good initial choice is a critical step, and we propose a successful method described further below.

3.1. The Generalized Least-Square approach (TVGLS) with positivity

A generalized linear admittance model including a normalized distribution $g_{\tau}(\tau)$ and additionally the so-called “high frequency dielectric response” (HFDR) or “capacitive response” $i\omega C$ (often effective in the data over 1 kHz), leads to consider one of the general linear expressions for impedance Z and admittance A spectrum as:

$$Z_{\omega}(\omega) = a_{\mathcal{H}} + M \int_0^{\infty} g_{\tau}(\tau) \varphi(\omega, \tau) d\tau, \quad (7)$$

or alternatively

$$A_{\omega}(\omega) = a_{\mathcal{H}} - M \int_0^{\infty} g_{\tau}(\tau) \varphi(\omega, \tau) d\tau + i\omega C. \quad (8)$$

In these expressions, $a_{\mathcal{H}}$ is a real constant (generally speaking σ_{∞} when using admittance), M is a real constant indicative of the amplitude of the IP proper phenomenon, $g_{\tau}(\tau)$ is the Relaxation Time Distribution (RTD), $\varphi(\omega, \tau)$ is the Kernel (for instance an elementary Debye response), C is a constant having the dimension of a capacitance and used to model the high frequency trend of the spectrum. Here the RTD, g is normalized ($\int_0^{\infty} g_{\tau}(\tau) d\tau = 1$).

A discussion about the sign of the integral factor is useful. Here the writing is chosen in such a way that we can assume that $M > 0$ in both cases. Indeed, if we consider the following elementary model:

$$\sigma^* = \sigma_{\infty} + \frac{\sigma_0 - \sigma_{\infty}}{1 + i\omega\tau_0} + i\omega C. \quad (9)$$

which is a special case of Eq. (8) with:

$$\begin{cases} a_{\mathcal{H}} = \sigma_{\infty} \\ M = -(\sigma_0 - \sigma_{\infty}) \geq 0 \\ g_{\tau}(\tau) = \delta(\tau_0 - \tau) \\ \varphi(\omega, \tau) = 1/(1 + i\omega\tau) \end{cases} \quad (10)$$

This model is compatible with the fact that $\sigma_0 - \sigma_{\infty} \leq 0$ and $|\sigma^*| \nearrow$ when $\omega \nearrow$.

Taking the inverse of the conductivity, the complex resistivity is given by:

$$\rho^* = \rho_{\infty} + \frac{\rho_0 - \rho_{\infty}}{1 + i\omega\tau_0}, \quad (11)$$

with:

$$\begin{cases} a_{\mathcal{H}} = \rho_{\infty} \\ M = \rho_0 - \rho_{\infty} \geq 0 \\ g_{\tau}(\tau) = \delta(\tau_0 - \tau) \\ \varphi(\omega, \tau) = 1/(1 + i\omega\tau) \end{cases}, \quad (12)$$

In terms of complex resistivity, this writing is compatible with fact that $\rho_0 - \rho_{\infty} \geq 0$ and $|\rho^*| \searrow$ when $\omega \nearrow$.

Note that in Eq. (9), the presence of the dielectric term is additive to the complex conductivity response assuming homogeneity. This choice

is actually imposed by the form of Ampere's law in which the dielectric displacement current is added to the conduction current. Therefore, the use of the following model for the complex resistivity:

$$\rho^* = \rho_\infty + \frac{\rho_0 - \rho_\infty}{1 + i\omega\tau_0} + i\omega C \quad (13)$$

would not be appropriate. Actually, in the presence of the HFDR, complex resistivity data have just to be first converted into complex conductivity data before being inverted in order to take advantage of the additivity of the HFDR.

Although the HFDR or “capacitive response” has been addressed recently by several authors (e.g., in Appendix 1 of Florsch et al., 2012), its origin (and order of magnitude) is still controversial. A recent advance can be found in Revil (2013). Although this paper involves porous media, it applies as well in homogeneous pure and salted water. In the following, we use in Eq. (8) the sign minus, that is we assume that the data to be inverted are relative to complex conductivity (or admittance) and we believe that this choice is actually dictated by the underlying physics of the problem discussed in our previous papers (e.g., Revil and Florsch, 2010).

3.1.1. Encapsulating M into $g\tau(\tau)$

In managing Eq. (7) or (8), one has to determine both the constant M and the RTD $g(\tau)$ with the additional condition that $g(\tau)$ must be normalized. This last condition could be undertaken by introducing an additional Lagrange multiplier to ensure this condition is obeyed. It is however simpler to encapsulate M into $g\tau(\tau)$, that is to write the model under the form (example with the minus sign):

$$A_\omega(\omega) = a_{\mathcal{R}} - \int_0^\infty g\tau(\tau)\varphi(\omega, \tau)d\tau + i\omega C, \quad (14)$$

by dropping both M and the normalization. Using this formalism, only $g\tau(\tau)$ has to be retrieved, and will not be normalized. However, it is trivial to recover the constant and the normalization condition just by integrating $g\tau(\tau)$ as a final step after the inversion of the RTD. This writing (making M disappearing from the inverse problem) is the real reason why we distinguish the two signs. In the following, we consider that the multiplicative constant has been therefore encapsulated in the RTD, and that a recovery of the true distribution requires normalization.

3.1.2. Using the logarithmic change of variables

We perform now the usual change of variables, which better accommodates the frequency (or time) scale invariance:

$$\begin{cases} z = -\log(\omega) \Leftrightarrow \omega = e^{-z} \\ s = \log(\tau) \Leftrightarrow \tau = e^s \end{cases} \quad (15)$$

This yields:

$$A_z(z) = a_{\mathcal{R}} - \int_{-\infty}^\infty G_s(s)\Phi(z, s)ds + ie^{-z}C, \quad (16)$$

$$\begin{cases} G_s(s) = \tau g(\tau) \\ \Phi(z, s) = \varphi(e^{-z}, e^s) \end{cases} \quad (17)$$

In most cases, $\varphi(\omega, \tau)$ is multiplicative, that is $\varphi(\omega, \tau)$ takes the form $\varphi(\omega, \tau)$ and then $\Phi(z, s)$ can be rewritten as $\Phi(z - s)$. It is not an absolute requirement. Actually using $\Phi(z, s)$ instead of $\Phi(z - s)$ is only useful if the “system” as properties varying with time. If Eq. (16) can be written as a convolution product and when excluding the dielectric term, the kernel can be retrieved by using a Fourier technique (see Florsch et al., 2012) or using classical fitting methods.

As an example, we consider the Havriliak–Negami (also called the generalized Cole–Cole model, see Havriliak and Negami, 1966; Pelton et al., 1983). This model is given by:

$$H_\omega(\omega) = a + \frac{d}{(1 + (i\omega\tau_0)^c)^b}, \quad (18)$$

(with $d < 0$ for admittance). The theory above shows that one can find a distribution relative to this model, of the form:

$$H_\omega(\omega) = a - \int_{-\infty}^\infty G_s(s)H_z(z-s)ds. \quad (19)$$

In this model, if we set $b = 1$, one deals with the Cole–Cole model, and then it can be written by applying the superposition principle with a simple Debye kernel, through the integral:

$$H_{\omega, b=1}(\omega) = a - \int_{-\infty}^\infty \frac{C_z(s)}{1 + ie^{s-z}}ds, \quad (20)$$

where $C_z(s)$ is the Cole–Cole distribution of relaxation time.

One may want to generalize this formalism based on Debey relaxation by opening to the possibility of fitting a spectrum with a superposition of (for instance) Havriliak–Negami models, that is, by operating a fit of a distribution $g\tau(\tau)$ in:

$$H_\omega(\omega) = a_{\mathcal{R}} - \int_0^\infty \frac{g\tau(\tau)}{(1 + (i\omega\tau)^c)^b}d\tau. \quad (21)$$

In that case, the problem is not anymore to retrieve *simultaneously* b , c and $g\tau(\tau)$, but to retrieve $g\tau(\tau)$ assuming b and c are known. Indeed, some physical consideration may lead geophysicists to tackle with such models. For instance one may want to decompose the data by using a superimposition of Warburgs' elementary responses instead of using a superimposition based on the Debye model. The Warburg model coincides with the Cole–Cole when $c = 0.5$ and $b = 1$. Although ancient (Warburg, 1899), this model has renewed interest because it permits to model properly the diffusion of ions in charge accumulations (Revil et al., 2014; Wong, 1979).

Remember that taking into account an additive HFDR requires the admittance formalism. This term is not directly additive if working with impedance, because the equivalent electrical circuit would be the impedance of two components in parallel, which are only harmonically additive. Additionally, Florsch et al. (2012) showed that it is not possible to encapsulate the dielectric term into the integral kernel. In other words, there is no bounded distributions that can be used to represent the dielectric term.

3.2. Inversion skeleton

The linear formalism, although we dropped it, is useful as a reference case, especially to arrange the covariance matrices and damping operators. Without loss of generality, the convolutive form is convenient for discussion, and we consider first the following equation:

$$A_z(z) = a_{\mathcal{R}} - \int_{-\infty}^\infty G_s(s)\Phi(z-s)ds + ie^{-z}C. \quad (22)$$

To model the unknown function G , digitizing (sampling) it is the simplest way (alternative method would for instance write G on a basis of continuous functions). The integral is replaced by its corresponding Riemann sum. For facility, one can split the spectrum into real and imaginary parts (although most of the computer languages allow the use of complex variables). Then we write Eq. (22) in the

new following form (with N_D the number of data, N_G the chosen number of samples of G and Δs the integration step):

$$\begin{cases} A_k^{\Re} = a_{\Re} - \Delta s \sum_{j=1}^{N_G} \Phi_{kj}^{\Re} G_j + \text{err}_k^{\Re} \\ A_k^{\Im} = -\Delta s \sum_{j=1}^{N_G} \Phi_{kj}^{\Im} G_j + C e^{-z_k} + \text{err}_k^{\Im} \end{cases} \quad (23)$$

In this equation, $A_{\omega}(\omega)$ has been separated in real and imaginary parts and irregularly sampled at the discrete frequencies ω_k . We also have: $G_j = G(s_j)$ and $\Phi_{kj}^{\Re} = \text{Re}(\Phi(z_k - s_j))$; $\Phi_{kj}^{\Im} = \text{Im}(\Phi(z_k - s_j))$ with Re and Im for the real and imaginary parts, respectively. The term (err) denotes the uncertainties in the measurements.

3.3. Setting the constant time range and exploration interval

To $G_j = G(s_j)$ correspond a set of $s_j = \log \tau_j$. For convenience, one uses from here the logarithm in base 10. One must set a time range $[\tau_{\min}, \tau_{\max}]$, a “working interval”, in which the unknown time distribution is sought. It is clear that we must involve the frequency interval boundaries in which the function to be inverted (the data) is known. Let us write $[f_{\min}, f_{\max}]$ this interval. A suitable choice is, in log:

$$\begin{cases} \log_{10}(\tau_{\min}) = -\log_{10}(2\pi f_{\max}) - N_1 \\ \log_{10}(\tau_{\max}) = -\log_{10}(2\pi f_{\min}) + N_2 \end{cases} \quad (24)$$

Actually we use, by experience, $N_1 = N_2 = 0$ but let us keep this notation to recall that the relationship between the frequency range and the time constant range is not absolute, and it is still useful in some case to make them varying. Using positive values means that the time range is greater than possibly derived from the frequency range. This leads to end effects, for instance at low time constant, the response is disturbed by the HFDR.

3.4. Setting the required matrices

The linear formalism remains useful to provide a first comprehensive and classical frame to the inversion, and to set the notations. To manage the linear inversion process, the system of equations (Eq. (17)) is better put in a matrix form:

$$\mathbf{P}\mathbf{x} = \mathbf{a} + \mathbf{Err}, \quad (25)$$

with:

$$\mathbf{P} = \Delta s \begin{pmatrix} \Phi_{11}^{\Re} & \Phi_{12}^{\Re} & \dots & \Phi_{1N_G}^{\Re} & 1 & 0 \\ \dots & \dots & \dots & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Phi_{N_D,1}^{\Re} & \dots & \dots & \Phi_{N_D,N_G}^{\Re} & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Phi_{11}^{\Im} & \dots & \dots & \Phi_{1N_G}^{\Im} & 0 & e^{-z_1} \\ \dots & \dots & \dots & \dots & \dots & e^{-z_2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Phi_{N_D,1}^{\Im} & \dots & \dots & \Phi_{N_D,N_G}^{\Im} & 0 & e^{-z_{N_D}} \end{pmatrix}; \quad (26a)$$

$$\mathbf{x} = \begin{pmatrix} G_1 \\ G_2 \\ \vdots \\ G_{N_G} \\ a_{\Re} \\ C \end{pmatrix}; \quad (26b)$$

and:

$$\mathbf{a} = \begin{pmatrix} A_1^{\Re} \\ A_2^{\Re} \\ \dots \\ A_{N_D}^{\Re} \\ A_1^{\Im} \\ \dots \\ A_{N_D}^{\Im} \end{pmatrix}. \quad (26c)$$

A classical least-square (LS) solution of this set is:

$$\hat{\mathbf{x}}_{\text{LS}} = (\mathbf{P}^t \mathbf{C}_{dd}^{-1} \mathbf{P})^{-1} \mathbf{P}^t \mathbf{C}_{dd}^{-1} \mathbf{a}, \quad (27)$$

where \mathbf{C}_{dd} is the diagonal covariance matrix of the data, and \mathbf{C}_{dd}^{-1} its inverse, namely: (28)

$$\mathbf{C}_{dd}^{-1} = \begin{pmatrix} (\sigma_1^{\Re})^{-2} & 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & (\sigma_{N_D}^{\Re})^{-2} & 0 & \dots & 0 \\ 0 & 0 & 0 & (\sigma_1^{\Im})^{-2} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & (\sigma_{N_D}^{\Im})^{-2} \end{pmatrix}. \quad (28)$$

3.5. Switching to the non-linear case

The system Eq. (23) or (25) is changed by setting that the unknown is now the exponential, that is:

$$\forall j, G_j = e^{G'_j}. \quad (29)$$

The system is now written with the unknown G'_j :

$$\begin{cases} A_k^{\Re} = a_{\Re} - \Delta s \sum_{j=1}^{N_G} \Phi_{kj}^{\Re} \cdot e^{G'_j} + \text{err}_k^{\Re} \\ A_k^{\Im} = -\Delta s \sum_{j=1}^{N_G} \Phi_{kj}^{\Im} \cdot e^{G'_j} + C e^{-z_k} + \text{err}_k^{\Im} \end{cases} \quad (30)$$

This set of equations seems similar to Eq. (25), but is however non-linear. One could be tempted to set $\forall j, G_j = e^{G'_j}$ and then solve it classically; but then we would give up positivity. Solving Eq. (25) and solving Eq. (30) are actually radically different. The positivity constraint is assured since even if some G'_j are found to be negative, $G_j = e^{G'_j}$ is always positive and never vanishes.

In the geophysical literature, the model vector is often noted \mathbf{m} (for model) or \mathbf{p} (for parameter), while the data vector is \mathbf{d} and the direct problem uses the notation \mathbf{G} , so the problem is written $\mathbf{d} = \mathbf{G}(\mathbf{m})$ (for instance in Menke, 1989). Here, $\mathbf{m} = (G'_1, G'_2, \dots, G'_{n_m})$ and $\mathbf{d} = (A_1^{\Re}, A_2^{\Re}, \dots, A_{n_d}^{\Re}, A_1^{\Im}, A_2^{\Im}, \dots, A_{n_d}^{\Im})$ and we use, $\mathbf{d} = \mathbf{f}(\mathbf{m})$, where \mathbf{f} is here a non-linear function of the unknown vector \mathbf{m} .

Turning back to the original least-square strategy (or formalism) is certainly a good way when dealing with the non-linear case. It consists

in posing the problem in terms of minimization of the weighted objective function, namely:

$$J_0(\mathbf{m}) = (\mathbf{d} - \mathbf{\Gamma}(\mathbf{m}))^t \mathbf{C}_{dd}^{-1} (\mathbf{d} - \mathbf{\Gamma}(\mathbf{m})) \quad (31)$$

Basically, it is necessary, but not sufficient nor satisfactory in general to get a minimum just solving $\nabla J_0 \equiv 0$. Actually, the set of solution verifying this condition, if it permits to fit the data, can correspond to unrealistic results (for instance, solutions showing huge oscillations). It is necessary to include a damping add-on (or regularizer), which is equivalent to introducing a priori knowledge on the solution. This is essential to restrict the wide solution space verifying Eq. (31) to pragmatic solutions, although not fully realistic (only a full Bayesian approach can avoid this difficulty).

3.6. Stable inversion and setting the damping parameters and a priori

The solution of the “ill-posed problem”, in Tikhonov’s sense (Tikhonov and Arsenin, 1977) can be “regularized”, or “damped” and this is done, in the frame of the optimization theory, by adding a damping (or penalization) function to the initial objective (or “cost”, or “misfit”) function.

The new complete damped objective function is:

$$J(\mathbf{m}) = J_0(\mathbf{m}) + J_1(\mathbf{m}), \quad (32)$$

that is, we aim to minimize J instead of J_0 only. One of the simplest choices to stabilize the solution is just taking a minimum norm regularizer:

$$J_1(\mathbf{m}) = \lambda^2 \|\mathbf{m} - \mathbf{m}^{\text{prior}}\|^2. \quad (33)$$

It is obvious that if λ is very small, the optimum will not change significantly, while if λ is big, J_1 is no more negligible with respect to J_0 and will attract the solution toward a given a priori parameter vector $\mathbf{m}^{\text{prior}}$ to minimize $\|\mathbf{m} - \mathbf{m}^{\text{prior}}\|^2$. (Squaring λ is just a writing to have this coefficient positive.)

In the following we use a more complex damping operator, namely:

$$J_1(G) = \lambda^2 \sum_{k=1}^{N_g-1} (G'_{k+1} - G'_k)^2 + \alpha^2 (a_{91} - a^{\text{prior}})^2 + \beta^2 C^2. \quad (34)$$

When acting on λ we purely smooth more or less the solution, limiting the norm of its second derivative. Finally, α and β permit to limit the excursion of $(a_{91} - a^{\text{prior}})$ and C , respectively, if wished.

Choosing the best damping parameters $\{\lambda_{\text{opt}}, \alpha_{\text{opt}}, \beta_{\text{opt}}\}$ can be a matter of trial-and-error and experience. It is problem-dependent, and, for a given problem, data dependent. It is always a question of trade-off, which must be adjusted at the end partially “by hand”. When chosen too small, the solution is unstable. When chosen too big, the solution leads to unsatisfactory fittings and the solution is flattened to zero. Because it is dependent on the choice of the operator, it is also a subjective part of the inversion, or, equivalently, the prior information embedded into the inversion scheme by the operator itself.

Most of the time the solution is satisfactory by using $\alpha = \beta = 0$. But if we expect that, for instance, the end of the spectrum is likely not due to a HFDR, then one may set β to a large value in order to limit the excursion of C , or just because we would like to cancel this term in the model. Finally, the main smoothing coefficient is λ . A very suitable method to set it is the so-called method of the L-curve. The L-curve method has been used successfully by Florsch et al. (2012) to invert RTD by using Fourier transform, and is detailed in Argavai (2003) and Hansen and O’Leary (1993).

3.7. The L-curve

The L-curve is the curve obtained by plotting the solution modulus $\|\mathbf{G}_\lambda\|$ depending on λ^2 versus the misfit $\|A_{\text{observed}}(\omega) - \tilde{A}_\lambda(\omega)\|$, when the damping parameter (here λ^2) is varying from a high value—typically 10^6 in our experience (this flattens the solution, that is $\|\mathbf{G}_\lambda\| \rightarrow 0$) to a small value, like 10^{-2} (this lets the solution to be unstable where $\|\mathbf{G}_\lambda\| \rightarrow \infty$). To run the method it is safe to prevent from overflow or underflow that may break the computation (i.e., the run of the algorithm should continue even if underflows and overflows occur). Numerical overflow may lead $\|\mathbf{G}_\lambda\|$ to be huge and unstable, and this happens when the solution is not dumped enough. Actually, this occurs because one benefit from exploring a λ range as wide as possible, to facilitate the L-curve corner location, where the damping is optimal.

To detect this optimal point, as depicted on Fig. 2, we use the following trick: the corner becomes the minimum of the curve once the latter has been rotated by an angle α counterclockwise.

This procedure proves efficient most of time while $\alpha \in [45^\circ, 80^\circ]$, but sometimes it can be tricky to find the “good” rotation, and in the example we show in this paper, we had to take one time $\alpha = 91^\circ$. However,

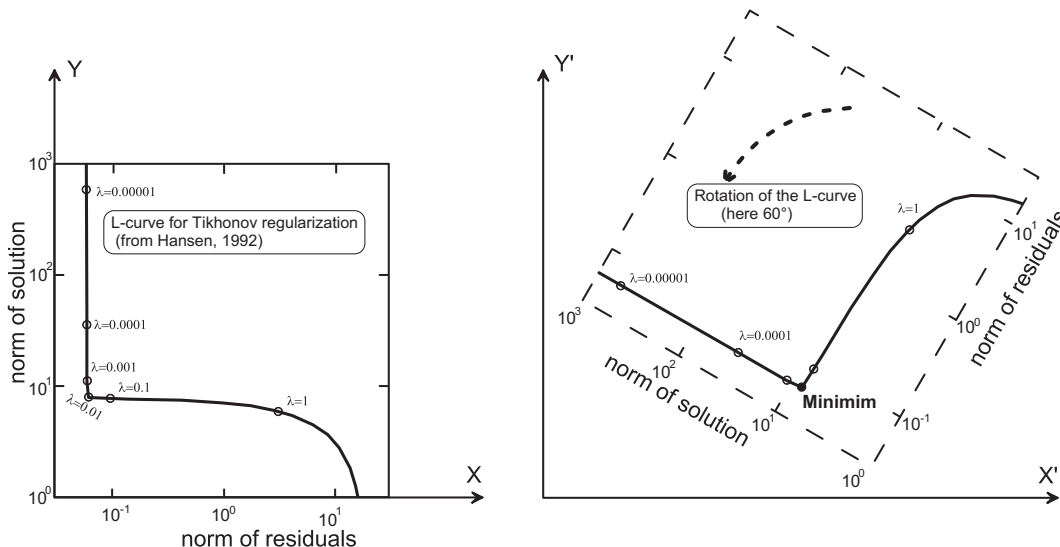


Fig. 2. Ideal L-curve valid in the linear case, and rotation of it, making the optimum damping parameter easily and automatically determined as a minimum.

this angle is the only parameter “managed by hand”, and the fact that the code fails for a representative value of saying $\alpha = 50^\circ$ may result from a L-curve badly shaped. The reason why the L-curve may be shaped very differently than in the paper by Hansen and O’Leary (1993) is a consequence of the non-linearity. For a linear problem, there is a unique solution (or a “valley” solution if not damped, but applying the L-curve makes the solution unique). That is, for a linear problem, the objective function is quadratic, a perfect paraboloid in N dimensions (N being the number of unknowns). Then, in the linear case, the L-curve looks really like an “L”. In the non linear case (due here to the positivity constraint we demand), local minimum emerges (see for instance Fernández-Martínez et al., 2012), making the choice of the optimum point on the L-curve trickier. In such a case, it is benefiting to check the data more in detail to see whether they are coherent or not with the kernel model, or whether the signal has not been disturbed by non IP contributions (like 50 or 60 Hz EM radiation, or low frequency electrode drifts).

3.8. Notice on possible correlation between $a_{\mathcal{H}}$ and C with $g_T(\tau)$

The parameter $a_{\mathcal{H}}$ is necessary in the formalism, and sometimes C is required (generally speaking when the range includes frequencies greater than 1 kHz) but both may be correlated with something of $g_T(\tau)$. Let us for instance consider, a decomposition on a basis of generalized Cole–Cole (or Havriliak–Negami) model and let us examine the two ends of the frequency range.

We have first:

$$A_\omega(\omega)|_{\omega=0} = \left[a_{\mathcal{H}} - \int_0^\infty \frac{g_T(\tau)}{(1 + (i\omega\tau)^c)^b} d\tau + i\omega C \right]_{\omega=0} \Rightarrow A_\omega(0) = a_{\mathcal{H}} - \int_0^\infty g_T(\tau) d\tau. \quad (35)$$

Clearly, a trade-off does exist between $a_{\mathcal{H}}$ and the integral $\int_0^\infty g_T(\tau) d\tau$, coming from this coupling at low frequency. It is partial, since the whole frequency interval influences $g_T(\tau)$.

Let us go a little bit further by expanding $A_\omega(\omega)$ close to $\omega = 0$. Since the distribution $g_T(\tau)$ cannot be explored over the whole time domain in practice, and especially not for very large time constant (because it will make the measurement duration prohibitive), we have no scruples in assuming that:

$$\exists T | g_T(\tau) = 0 \text{ if } \tau \geq T. \quad (36)$$

(In other terms: even if it may be physically unlikely, but anyway we cannot reach the corresponding time constant domain).

Then we can consider the frequency domain where ω is small enough to have $\omega T \leq \omega T \ll 1$, (that is $\omega \ll 1/T$) and we can write at the first order:

$$A_\omega(\omega) = a_{\mathcal{R}} - \int_0^T \frac{g_T(\tau)}{(1 + (i\omega\tau)^c)^b} d\tau + i\omega C \cong a_{\mathcal{R}} - \int_0^T g_T(\tau) [1 - (i\omega\tau)^{cb}] d\tau + i\omega C. \quad (37)$$

Introducing the moment (limited to T) of non integer order γ by:

$$m_\gamma = \int_0^T \tau^\gamma g_T(\tau) d\tau \quad (38)$$

We can now write:

$$A_\omega(\omega)|_{\omega T \ll 1} \cong a_{\mathcal{R}} - m_0 - (i\omega)^{cb} m_{cb} + i\omega C. \quad (39)$$

Separating real and imaginary parts in this equation leads to:

$$A_\omega(\omega)|_{\omega T \ll 1} \cong \left[a_{\mathcal{R}} - m_0 - m_{cb} \omega^{cb} \cos\left(\frac{\pi cb}{2}\right) \right] + i \left[-m_{cb} \omega^{cb} \sin\left(\frac{\pi cb}{2}\right) + \omega C \right]. \quad (40)$$

In the case of the simple Debye, we have:

$$A_\omega(\omega)|_{\omega T \ll 1} \cong (a_{\mathcal{R}} - m_0) + i\omega(-m_1 + C) \quad (41)$$

Generally speaking, the HFDR is only seen at the high frequency end, generally over 100 Hz or even 1000 Hz. Then it is likely that $|m_1| > |C|$, while this is not resulting from mathematics, but experimentation. Then it can even be simplified to:

$$A_\omega(\omega)|_{\omega T \ll 1} \cong (a_{\mathcal{R}} - m_0) - i\omega m_1 \quad (42)$$

This equation shows the content of the left, low-frequency, end of the spectrum.

At the opposite of the spectrum (at high frequency) the response is not bounded:

$$A_\omega(\omega)|_{\omega \rightarrow +\infty} = \left[a_{\mathcal{H}} - \int_0^\infty \frac{g_T(\tau)}{(1 + (i\omega\tau)^c)^b} d\tau + i\omega C \right]_{\omega \rightarrow +\infty} \cong a_{\mathcal{H}} + i\omega C. \quad (43)$$

It results in:

$$\begin{cases} \lim_{\omega \rightarrow \infty} \text{Re}[A_\omega(\omega)] \equiv a_{\mathcal{H}} \\ \lim_{\omega \rightarrow \infty} \text{Im}\left[\frac{A_\omega(\omega)}{\omega}\right] \equiv C \quad \text{or} \quad \lim_{\omega \rightarrow \infty} [\text{Im}(A_\omega(\omega))]_{\omega}' = C \end{cases} \quad (44)$$

This can be used to obtain first insights on the parameters (however the full determination requires the full inversion), or it can be advantageously be used as a priori values for $a_{\mathcal{H}}$ and C . Additionally we notice that since the time distribution appears no more in these equations, it results that the decoupling of $a_{\mathcal{H}}$ and C with respect to $g_T(\tau)$ is “provided” thanks to the upper part of the spectrum. In other words, the high frequency content brings almost no information to assess $g_T(\tau)$: it is dominated by these two constant contributions, $a_{\mathcal{H}}$ for the real part, and C for the imaginary one.

3.9. Initial guess and priors

It is useful and often necessary to launch the inversion from a realistic initial estimation, and in the TVGLS (Tarantola and Valette Generalized Least Squares), the initial guess is also the prior. The result just above can be used, but provides no insight about the RTD. Then our computation consists in beginning the optimization by departing the iteration scheme from an initial guess relative to what we call below the “homogeneous case”.

This homogeneous case corresponds to a model in which the distribution is assumed to be constant on a given interval, that is:

$$G_s(s) = \tau g_T(\tau) = \text{constant} = G_0 \quad (45)$$

over the given interval $\tau \in [T_{\min}, T_{\max}]$.

It results that $g_T(\tau) = \frac{G_0}{\tau}$ and it follows a model of the form:

$$\sigma^* = \sigma_\infty - \int_{T_{\min}}^{T_{\max}} \frac{1}{\tau} \frac{G_0}{(1 + i\omega\tau)^b} d\tau + i\omega C. \quad (46)$$

It is worthwhile to notice that the distribution $g_T(\tau) = G_0/\tau$ is the one involved in the notion of « Jeffrey » parameters, what Tarantola and Valette (1982) named initially the “null information” and later the “homogeneous probability measure” (Tarantola, 2005).

Eq. (46) can be easily integrated. It is also given, for instance, by Barsoukov and MacDonald, (2005, pp. 38–39). We write here the solution in the following explicit form with real and imaginary parts separated:

$$\begin{cases} R(\omega) = \sigma_{\infty} - G_1 \left[\log \left(\frac{T_{\max}}{\sqrt{1 + \omega^2 T_{\max}^2}} \right) - \log \left(\frac{T_{\min}}{\sqrt{1 + \omega^2 T_{\min}^2}} \right) \right] \\ I(\omega) = -G_1 \left[\tan^{-1}(\omega T_{\min}) - \tan^{-1}(\omega T_{\max}) \right] + iC\omega \end{cases} \quad (47)$$

For the support time interval, we use the simple choice:

$$W_T = [T_{\min}, T_{\max}] = \left[\frac{1}{2\pi f_{\max}}, \frac{1}{2\pi f_{\min}} \right]. \quad (48)$$

Replacing angular velocities by the sampled values at discrete frequencies, Eq. (47) leads to a system of linear equations with robust unknowns $\{\sigma_{\infty}, G_1, C\}$ which can be solved by a classical least-square approach. In the inversion code, the solution is taken as initial and prior values, and especially the unknowns relative to the RTD to be inverted are set to the constant G_0 . The little code required to get this initial model by inversion is a part of our code, used just before the launch of the main inversion procedure.

3.10. The damping and its connection with a priori covariance matrix

The Bayesian approach offers a very general background to inversion philosophy. In this framework, regularization methods are visited within the probabilistic theory by considering equivalent model a priori probability density function (or probability measures if not normalizable). The books by Tarantola (2005) and Menke (1989) both provide a complete and useful background for any reader interested in deepening the inverse problem methods. A demonstrative application of the Bayesian approach as applied to IP can be found in Ghorbani et al. (2007) in which this approach was applied to the Cole–Cole model.

Assuming the data errors are Gaussian, the Bayesian approach leads to consider the following probability density function, (with \mathbf{m}_0 an a priori on the model vector \mathbf{m} , which also plays the role of an attractor for \mathbf{m} toward \mathbf{m}_0):

$$\text{pdf}(\mathbf{m}) = K \exp \left\{ -\frac{1}{2} \left[(\mathbf{d} - \Gamma(\mathbf{m}))^t \mathbf{C}_{dd}^{-1} (\mathbf{d} - \Gamma(\mathbf{m})) + (\mathbf{m} - \mathbf{m}_0)^t \mathbf{C}_{mm}^{-1} (\mathbf{m} - \mathbf{m}_0) \right] \right\}. \quad (49)$$

Looking for the maximum of the pdf is then equivalent with looking to the least-square solution by identifying (see Eqs. (32), (33) and (34)):

$$J_1 \equiv (\mathbf{m} - \mathbf{m}_0)^t \mathbf{C}_{mm}^{-1} (\mathbf{m} - \mathbf{m}_0). \quad (50)$$

The problem is then to minimize:

$$J = J_0 + J_1 \equiv (\mathbf{d} - \Gamma(\mathbf{m}))^t \mathbf{C}_{dd}^{-1} (\mathbf{d} - \Gamma(\mathbf{m})) + (\mathbf{m} - \mathbf{m}_0)^t \mathbf{C}_{mm}^{-1} (\mathbf{m} - \mathbf{m}_0). \quad (51)$$

In correspondence with the term of Eq. (33), the damping contribution takes the more complete form:

$$J_1(G) = \lambda^2 \sum_{k=1}^{N_g-1} (G'_{k+1} - G'_k)^2 + \alpha^2 (a_{\mathcal{N}} - a^{\text{prior}})^2 + \beta^2 C^2. \quad (52)$$

Here, the inverse of the new covariance matrix relative to the model is:

$$\mathbf{C}_{mm}^{-1} = \lambda^2 \mathbf{W}_1 + \alpha^2 \mathbf{W}_2 + \beta^2 \mathbf{W}_3, \quad (53)$$

with:

$$\mathbf{W}_1 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 & 0 \\ \dots & & & \dots & -1 & 0 \\ & & 0 & -1 & 1 & 0 \\ 0 & & & & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 \end{pmatrix}; \quad (54a)$$

$$\mathbf{W}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & & & \dots & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ 0 & & & 0 & 1 & 0 \\ & & & 0 & 0 & 0 \end{pmatrix}; \quad (54b)$$

and

$$\mathbf{W}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & & & \dots & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ 0 & & & 0 & 0 & 0 \\ & & & 0 & 0 & 1 \end{pmatrix}. \quad (54c)$$

3.11. Posterior uncertainty assessment

We state below that for the class of problem we are dealing with, the uncertainty estimates are not meaningful when damping is used. Generally speaking, when performing an inversion, it is generally useful (sometimes mandatory) to provide uncertainties and error bars for the solutions. Within the linear inversion framework (which can be used in the non linear case as well), the information associated with the data is given by the posterior covariance matrix, the data resolution matrix, and the model resolution matrix (Menke 1989).

The posterior covariance matrix is nothing but the resulting uncertainties of the inverted parameters and the “correlation” between them. It is symmetrical. The diagonal of the matrix is directly related to the variance of the model parameters.

However and despite a linear formalism (extendable to the non-linear case), our conviction is that estimating the final uncertainties can be strongly misleading. For simplicity, we discuss the linear case, in which the L-curve looks like a L. It is easy to understand what happens when considering the vertical branch of the L-curve (before rotation). Within the vertical branch, the fit is reached and fully satisfactory in the data space. However, it corresponds to unstable solutions, even divergent. Mathematically speaking, those solutions are possible and even good-looking in terms of data fit. Then they should be taken into account in the estimator standard deviation. This makes nonsense, because satisfying solution (in terms of fit) may be definitely huge: then uncertainty assessment is not relevant. Moreover, if we consider the optimum point, the formal error is depending on the value of the damping

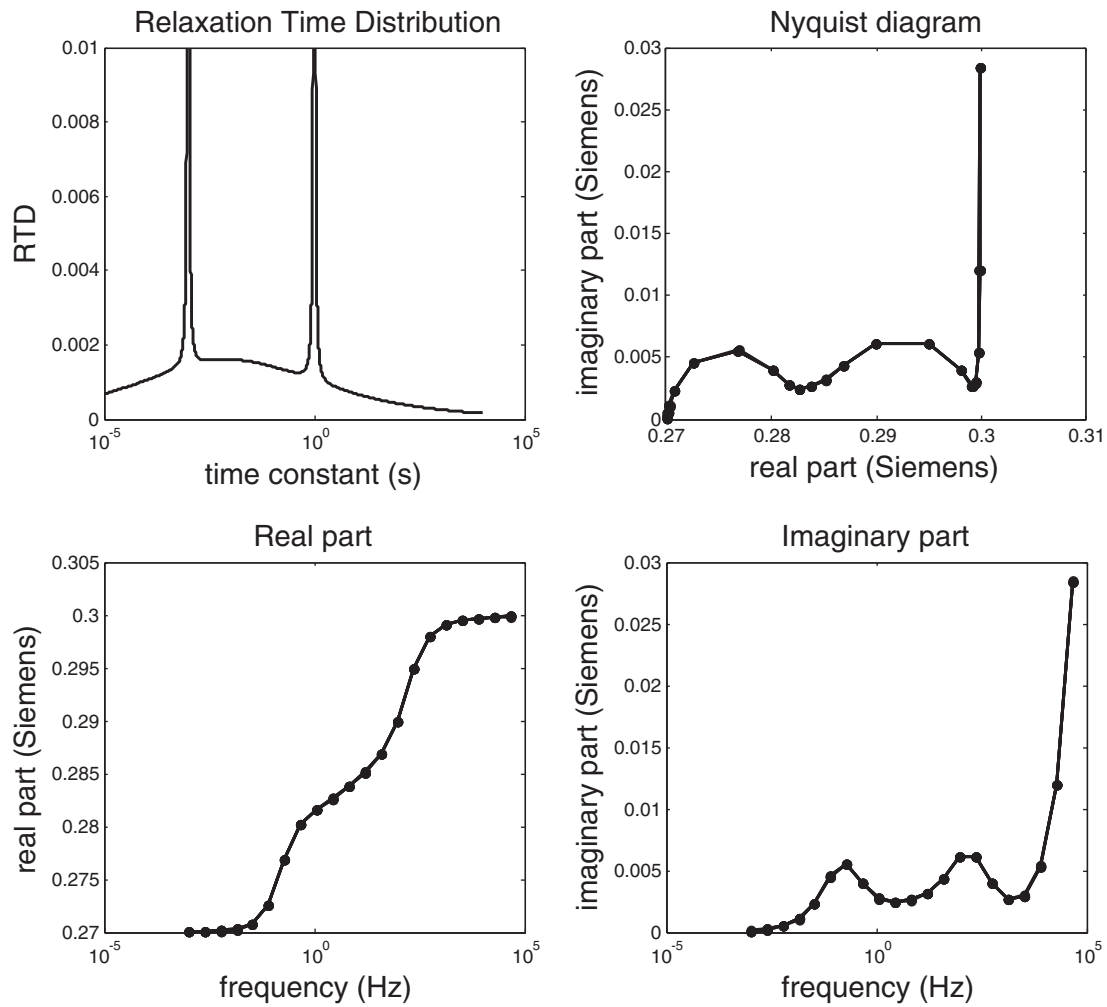


Fig. 3. Synthetic case. The RTD is made of two Dirac and one Warburg distributions. An HF dielectric term is added to the spectrum, shown in both Bode and Nyquist diagrams.

parameter, more than on the data themselves. Hence, we believe that providing the posteriori covariance matrix is not relevant because of its dependence on the damping parameter. Posteriori statistics is only significant if no damping is required. One must therefore be very careful regarding uncertainties estimates when dealing with ill-posed problems.

3.12. Tarantola–Valette Generalized Least-square inversion scheme

Eq. (23) (or 24 or 25) from Tarantola and Valette's (1982) paper is easy to implement and provides an iterative scheme with Newton–Raphson convergence properties. More precisely, the iterations are performed for a given fixed L-curve parameter (λ^2). That said, in an

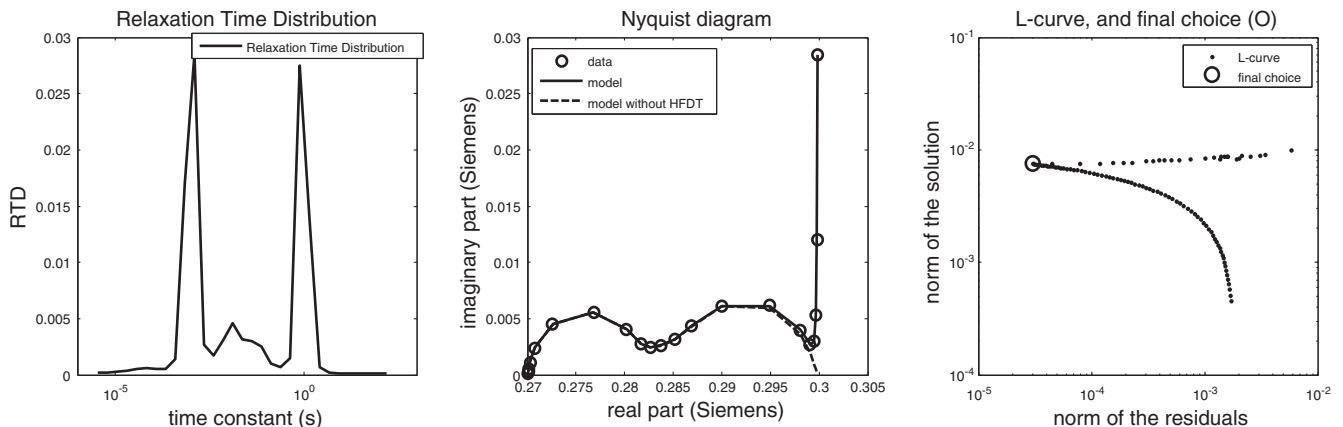


Fig. 4. Inversion result of the synthetic model by using a Debye decomposition.

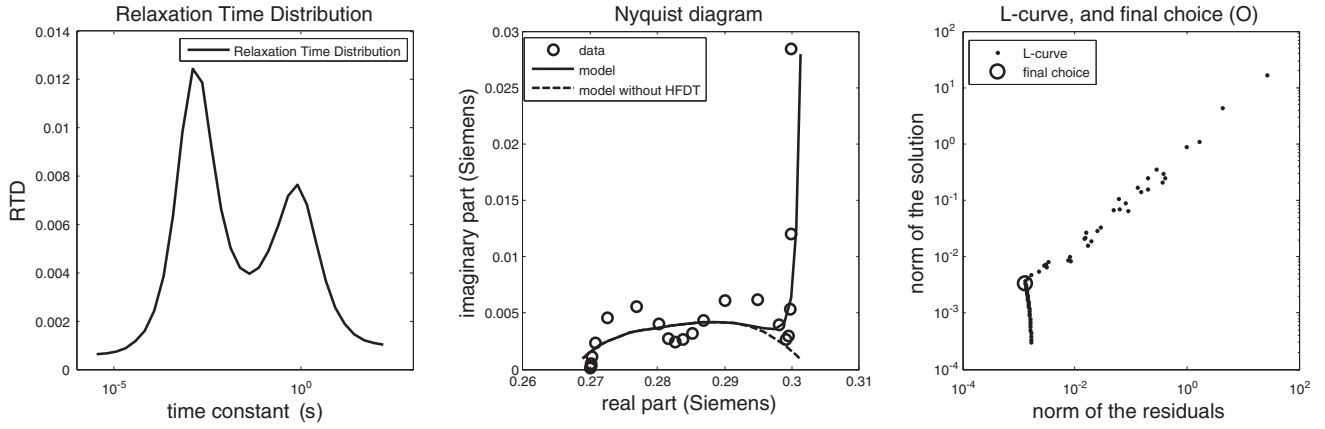


Fig. 5. We try to represent a spectrum involving two Dirac distributions in the RTD on the basis of a Warburg function, and as expected, it does not work.

external additional loop, (λ^2) is swept along within an interval $[\lambda_1, \dots, \lambda_2, \dots, \lambda_{N_\lambda}]$, browsing a given logarithmically-sampled interval.

4. Application to the Pore Size Distribution (PSD)

A general relationship between the relaxation time and a polarization length scale (generally the particle size) is consistent with many published works since the seminal work on particle size by Schwarz (1962). Recently, this idea has been generalized by Revil et al. (2012, 2014) to pore throat size distributions. This relationship connects the relaxation time constant to a representative size of the entity carrying the polarization process and called the polarization length scale. The general relationship has the form:

$$\tau = \frac{r^2}{kD}, \quad (55)$$

where r denotes a characteristic size (like pore throat size or particle radius), k is a constant (e.g., $k = 2$ in the work of Schwarz, 1962), and D is the diffusion coefficient of the charge carriers.

It results in a linear relationship between the logarithms of r and τ which makes easy the transformation of the respective distributions. The distribution shape is therefore invariant, just involving a translation along the x -axis (depending on a factor $K = kD$) and a scaling factor $1/2$. Then, the GSD (for Grain Size Distribution) or the PSD (for Pore Size Distribution) are similar to the RTD (Relaxation Size Distribution), just by changing the x -axis linear graduation.

5. Results: inversion code call and examples

The code available as a supplementary file is written in Matlab and the call uses and syntax are detailed in the Appendix A.

To illustrate the code, we apply these equations to one synthetic and two different cases of real data sets. Here we use three projection bases: i) Debye ($c = b = 1$), ii) Warburg (equivalent to a Cole–Cole with $c = 0.5$, $b = 1$) and iii) generalized Cole–Cole, with $c = 1$ and $b = 0.5$, which is also a special case of the Davidson–Cole model. Of course any desirable c and b may be used, within the range $]0,1[\times]0,1[$.

5.1. A synthetic (noisy) example

We first propose a synthetic model built with a sum of two Dirac distributions and a Warburg distribution. Actually we approximate the two Dirac distributions by using a Cole–Cole with $c_1 = c_3 = 0.999$, and the Warburg is nothing but a Cole–Cole with $c_2 = 0.5$. The two Dirac distributions are set at periods of $\tau_1 = 1$ s; $\tau_3 = 1$ ms, while the Warburg has a central period at $\tau_2 = 10$ ms. All models have an amplitude of $A_1 = A_2 = A_3 = 0.01$ (arbitrary units since it is synthetic), the level of random Gaussian noise standard deviation has been set at 10^{-5} with 0 mean, noted $n(0, 10^{-5})$. Finally, the HFDE is made up from a equivalent capacitance of $0.1 \mu\text{F}$ ($C = 10^{-7}$) and the additional term is $a = 0.3$. Then the synthetic model is shown on Fig. 3, and corresponds to the formula:

$$S(\omega) = a + \sum_{k=1}^3 \frac{A_k}{1 + (i\omega\tau_k)^{c_k}} + i\omega C + n(0, 10^{-5}). \quad (56)$$

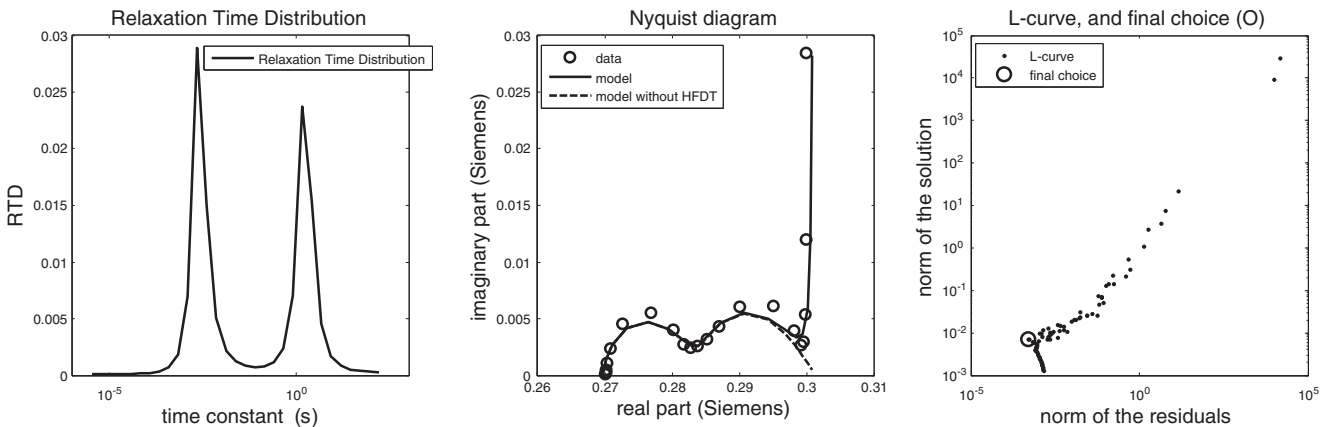


Fig. 6. Fit of the synthetic model with a Havriliak–Negami (or Generalized Cole–Cole with the particular case $c = 1$ and $b = 0.5$). The fit is not as good as when using Dirac function, but provides better results than fitting on a Warburg basis. The L-curve displays a random-like behavior in the vicinity of the optimum point.

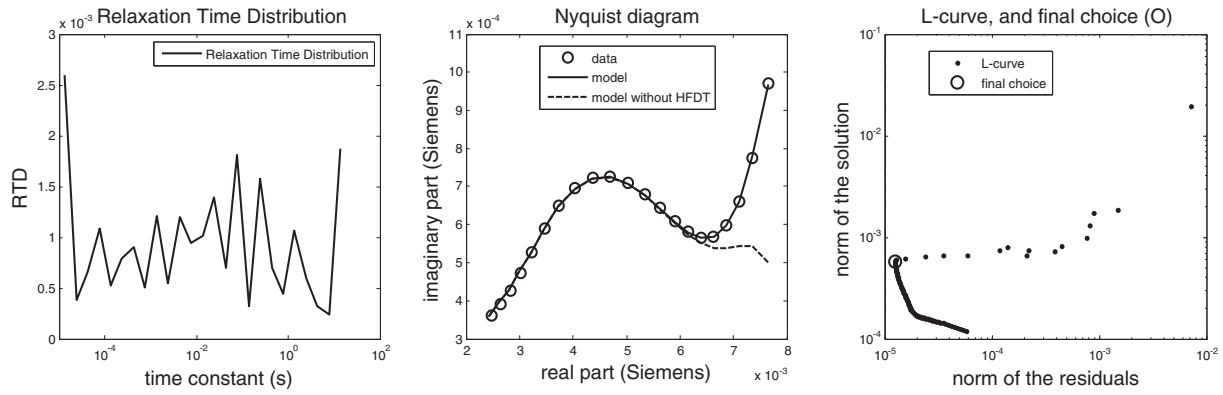


Fig. 7. Inversion of real data set from Aulus (Florsch et al., 2011), by using a Debye basis. Note that choosing the L-curve optimum is tricky. In the present case, it leads to oscillatory solutions, which are not satisfying. See also Fig. 8 which concerns the same model, but inverted by using another L-curve corner point.

The distribution relative to a Cole–Cole model with a time constant τ_0 and an exponent c is given by (Cole and Cole, 1941):

$$g(\tau) = \frac{1}{2\pi\tau} \frac{\sin(1-c)\pi}{\cosh\left[c \log\left(\frac{\tau}{\tau_0}\right)\right] - \cos(1-c)\pi}; \quad (57)$$

Then the analytical RTD to recover is just the sum of three functions of that kind with proper time constants, c constants and corresponding amplitudes, as defined above.

We provide three inversion cases, with the inversion using projection on three kinds of distributions:

- 1) on a Debye basis ($c = b = 1$ in Eq. (21)). See Fig. 4.

The call command used here is:

```
RTDINV('DWD.dat', 'RTD.dat', 'FIT.dat', 0, 1, 1, 50, 4e-12, 0, 0, 2, 4, 100, 89, 1, 1, 0, 12)
```

The recovery of the RTD is rather good, and the high frequency term correctly decoupled. As explained in the text, the L-curve ideal shape is lost due to the non linearity, but the optimum point is very well defined. Looking in detail, the Dirac distributions are not so accurate than in the synthetics, but it is not surprising when considering the resolution. At the contrary, the Warburg impedance contribution seems to be sharper than in the reality.

- 2) on a Warburg basis ($c = 0.5; b = 1$). See Fig. 5.

Knowing the solution, we can expect that the solution of the inverse problem will have some problems since it is not possible to project Dirac distributions on the continuous Warburg function. That said, of course, in real life, we possibly do not have such a prior knowledge

of the solution. Without surprise, the fit is bad; no RTD based on Warburg response may accommodate Dirac distributions. This example shows that before performing the inversion, the model should be realistic.

Command is:

```
RTDINV('DWD.dat', 'RTD.dat', 'FIT.dat', 0, 0.5, 1, 50, 4e-12, 0, 0, 2, 4, 100, 89, 1, 1, 0, 12)
```

- 3) on a Havriliak–Negami (or Generalized Cole–Cole), with the particular case ($c = 1; b = 0.5$). See Fig. 6.

We try to fit the synthetic spectrum by projecting over a Havriliak–Negami function basis. This is also supposed not to be good and indeed the results are poor (while not as poor as in the Warburg case). Only the two Dirac distributions are seen, and the fit is not as good as in the Dirac basis case.

Command:

```
RTDINV('DWD.dat', 'RTD.dat', 'FIT.dat', 0, 1, 0.5, 50, 4e-12, 0, 0, 2, 4, 100, 89, 1, 1, 0, 12)
```

One first conclusion that we can draw from this synthetics example is that the L-curve shape may contribute as good index to qualify the relevancy of the chosen model.

5.2. Example 2: data from archaeological slags

It is taken from iron slags that have been studied at an archaeological campaign (Florsch et al., 2011). The response of the concentrated slags is quite similar but stronger than the data acquired on the field (see Fig. 7 in the cited paper). The first inversion we propose is computed by using $\theta = 85^\circ$ when rotating the L-curve. Results are shown on Fig. 7.

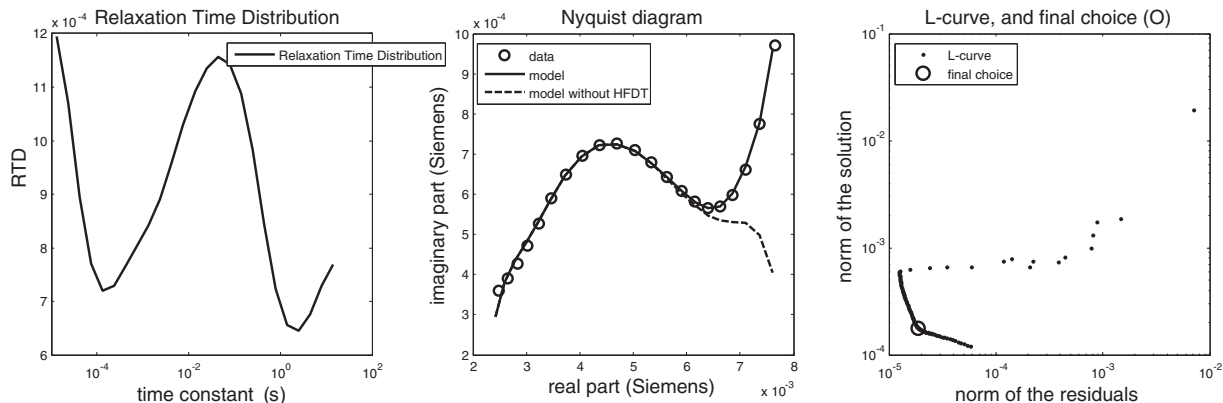


Fig. 8. Inversion of real data set from Aulus (Florsch et al., 2011) by using a Debye basis (see Fig. 7 for a comparison with another optimum point on the L-curve). Here, we reach another candidate point on the L-curve, and get smoother RTD, despite a bad fit at the low frequency end of the data.

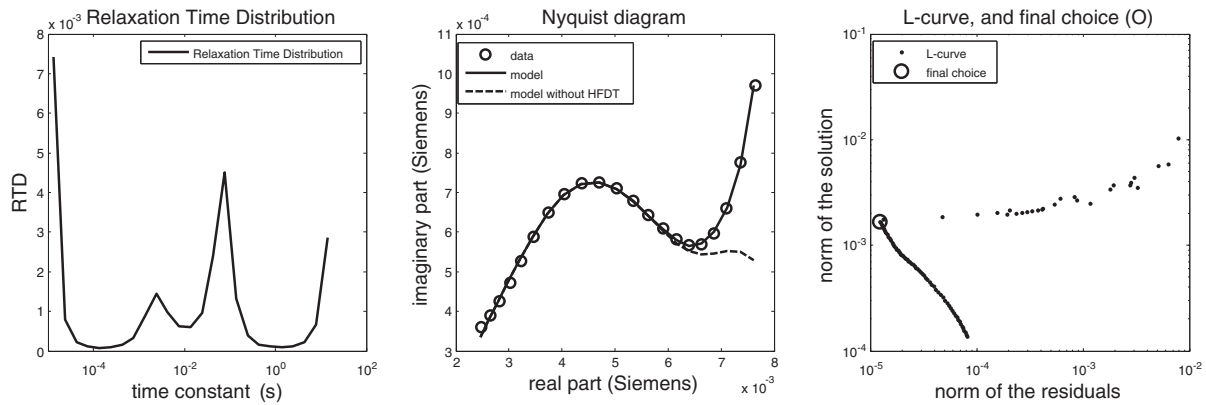


Fig. 9. Inversion of real data set from Aulus (Florsch et al., 2011), by using a Warburg basis, to be compared to Figs. 7 and 8. The shape of the L-curve makes the inversion reliable, while the results show that a single Warburg function would have been suitable to invert this data set since the deconvolved solution is close to a Dirac distribution.

Although the fit is excellent, the solution is oscillating and nonphysically realistic. In this case, the command is:

```
RTDINV('PYR.dat', 'RTD.dat', 'FIT.dat', 1, 1, 1, 50, 4e-12, 0, 0, 3, 4, 100, 85, 1, 1, 0, 12)
```

Then we wished to reach the less optimal L-curve which can be seen on the L-curve, as plotted on Fig. 8. To get this result, we set $\theta = 45^\circ$, using the command line:

```
RTDINV('PYR.dat', 'RTD.dat', 'FIT.dat', 1, 1, 1, 50, 4e-12, 0, 0, 3, 4, 100, 45, 1, 1, 0, 12)
```

The fit is less efficient in the previous case, especially at the low frequency limit. However, the solution is smoother. The lower end of the spectrum is not adjusted by the model, and actually both ends are questionable.

Let us consider now the use of a Warburg basis. Before inverting the data, we observe that the solution obtained on the Debye basis leads to a very smooth solution. Since Warburg function is much more spread than Debye function, we expect obtaining better results. The command is here:

```
RTDINV('PYR.dat', 'RTD.dat', 'FIT.dat', 1, 0, 5, 1, 50, 4e-12, 0, 0, 3, 4, 100, 85, 1, 1, 0, 12)
```

The solution shows that we were right (Fig. 9). The L-curve first is non ambiguous, making the geophysicist to be more confident with the inversion result. The data are well fitted, except possibly, like before, at the very lower frequencies (0.01 Hz). Apart the effects at the two ends, one main peak and one secondary peak are clearly defined, close

to 0.1 and 0.001 s respectively. At least the biggest one is significant, but the more remarkable here is that the data can be finely fit on a Warburg decomposition model, compatible with the so-called “electrode polarization” generally encountered with particle showing electronic conductivity. The fact is that the field results and some analysis make the magnetite the main candidate that may originate this signal.

5.3. Example 3: a mudstone from Wyoming

The sample is a mudstone originated from the Great Divide Basin in Wyoming, USA. Its permeability is only 1 mD (equivalently 10^{-15} m^2). Both Nyquist and Bode diagrams reveal outliers in the lower part of the spectrum. As previously, it appears below 1/100 Hz, and may indicate that the governing physic we assume here is not relevant, or at list not dominant below this spectrum. Taking into account the relative duration, that is more than 1 or 2 min, it could be linked with a long-distance drift of bulk ions. We do not develop here this phenomenon but just concentrate to the inversion. Then, to prevent spurious effects, we weighted the low part of the spectrum in such a way the corresponding parts do not impact the inversion.

An inversion by using Debye decomposition, as shown on Fig. 10, leads to quite satisfying fits. However, choosing the L-curve optimum is clearly problematic. We used an angle of 91° . If we had used a smaller angle (90° for instance), it is the lower point that would have been targeted, with still a good fit, but a totally flattened RTD.

The command is here:

```
RTDINV('439-1.dat', 'RTD.dat', 'FIT.dat', 1, 1, 1, 50, 4e-12, 0, 0, 5, 4, 100, 91, 1, 1, 0, 12)
```

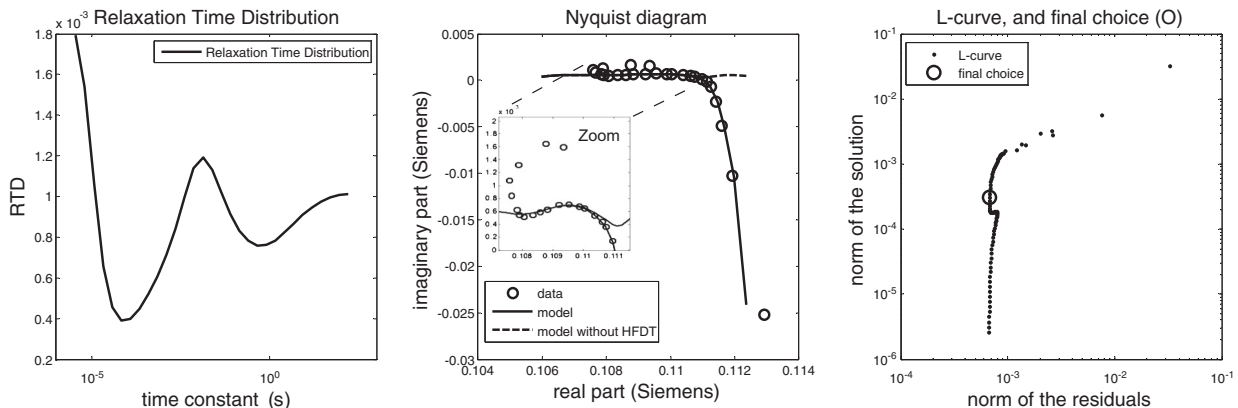


Fig. 10. Inversion of real data set relative to a mudstone of the Great Divide Basin (Wyoming, USA).

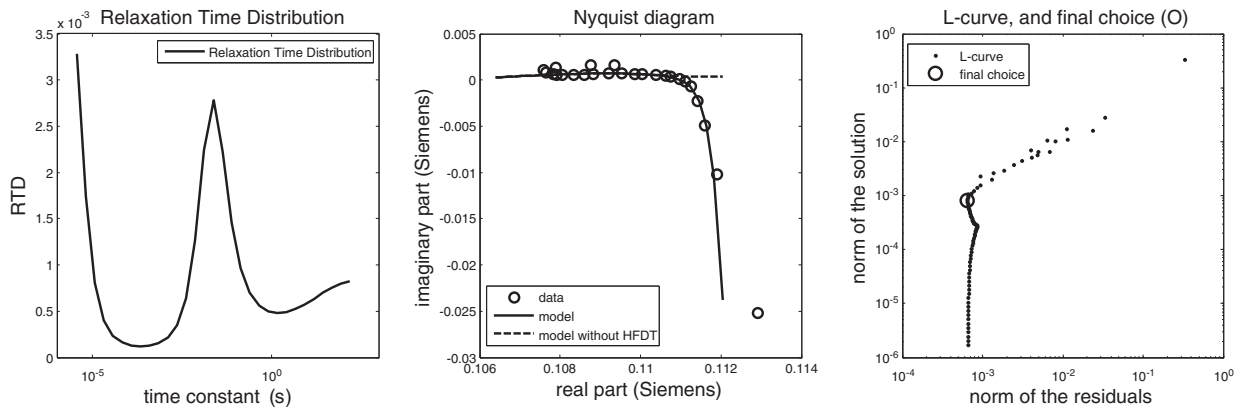


Fig. 11. The same as that in Fig. 10, but assuming a Warburg decomposition basis.

These data however can be fit smarter by using a Warburg basis (Fig. 11). The difficulty inherent to the L-curve optimum choice is still here, but the residual cost is the same, around 6.10^{-4} , and the angle used here is 90° , relatively to the command line:

```
RTDINV('439-1.dat','RTD.dat','FIT.dat',1,0.5,1, 50,4e-12,0,0,5,4,100,
91,1,1,0,12)
```

6. Conclusion

In induced polarization problems, the Green function of the problem is not always a Debye model and therefore spectral induced polarization spectra with Debye function to recover the distribution of relaxation time can be misleading. We offered in this paper a more general framework in which the operator can chose the Green function of the problem (base on underlying physics) and decompose the spectral induced polarization spectra using this Green function (e.g., a Warburg model). When inverting data and especially in the case of such ill-posed problems, the prior information is strongly involved in the solution, making the process depending on the operator. This makes the results shown in this paper depending on the strategy used, which include the following choices: (1) The mathematical model (Green function formalism with superposition principle), (2) The integration is made by using Riemann's summation, (3) A "High Frequency Dielectric Term" (HFDT) is included (but can be dropped when running the program), (4) The solution is seek in a given interval, and the operator decides what must be the range of the RTD involved, (5) The L-curve is used to provide an optimum damping parameter.

Hence the geophysicist may use and modify that Matlab program given as supplementary file (with data sets too) to get a better harmony with his expectation, but should never forget how the results depend strongly on our previous knowledge and overall prior knowledge. The use of the L-curve, associated with the Tarantola and Valette Generalized Least-Square, is one suitable method among other inversion methods, and leads to solution that are generally satisfying. However, it remains difficult to estimate uncertainties, and actually, the choice of the inversion method is more determinant on the solution than any method to propagate the uncertainties from the data space to the parameter space.

Acknowledgements

Nicolas Florsch thanks the ADEME-BRGM-IRD organizations through the contract no. 10006C0111-4491A1 for supporting this study. Christian Camerlynck and Nicolas Florsch also work in the context of the Equipex CRITEX. André Revil thanks Chevron Energy Technology Company (Grant # CW852844).

Appendix A. Inversion call syntax and format

The code is called by a Matlab command. It is open source and may be used, improved and distributed on a free basis.

The syntax of the command is given as follows:

```
RTDINV(name1,name2,name3,lecture_code,c,b,k,D,const_eps,
const_a,L_CENTRAL, L_RANGE,teta,factor,LI,ifig,size)
```

Parameters are described here below.

- Name1 is a string giving the name of the data file to be processed (only numeric data)
- Name2 is a string giving the name of the file devoted to RTD or SIZE DISTRIB
- Name3 is a string giving the name of the file devoted to the fitted spectra (amp, phase(mrad))
- lecture_code determine the reading format, with two choices:
 - if lecture=0, the format is
Freq(Hz), real_part (S/m), imaginary_part(S/m), err_real(S/m), err_imag(S/m)
(notation are straightforward, notice that all data concern conductivity and not resistivity)
 - if lecture = 1, the format is
Freq(Hz), amp (S/m), phase(mrad), err_amp(S/m), err_phase(mrad)
- c is the Cole–Cole exponent
- b is the exponent in the generalized Cole–Cole term $(1 + i\omega\tau^c)^b$
- const_eps is a damping parameter for the HFDT (β^2 in Eq. (34))
- const_a is a damping parameter for the additive constant (α^2 in Eq. (34))
- L_CENTRAL is the logarithm of the central value λ^2 around which λ^2 is swept
- L_RANGE is the half of the range regarding λ^2 , that is: $\log_{10}\lambda^2$ run through the interval:

$$([L_CENTRAL - L_RANGE, L_CENTRAL + L_RANGE])$$

- teta is the angle of rotation (in deg) in the L-curve diagram, to which the L-curve is rotated to facilitate the determination of the optimum λ^2
- factor is a multiplicative factor that may be applied to the errors. It equals 1 in general, but if we suspect the error to be unrealistically too small (example: only formal errors are given, but systematic noise is known to be present), then one can set factor > 1.
- LI is 0 or 1 and permits to choose to plot the L-curve on log-log or lin-lin
- ifig is 0, 1 or 2. This parameter is relative to three levels when plotting results. Level 0 is the basic graphics (identical to those given in this paper), 1 is intermediate and 2 provides more figures (including formal covariance).
- size determines the size of the labels used in the output figures

Appendix B. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jappgeo.2014.07.013>.

References

- Agarwal, V., 2003. Total variation regularization and L-curve method for the selection of regularization parameter. ECE 599, available on line at <http://imaging.utk.edu/people/former/vivek/EdgeTV.pdf>.
- Barsoukov, E., MacDonald, J.R., 2005. Impedance Spectroscopy Theory, Experiment, and Applications. edited by Second Edition John Wiley & Sons, Hoboken (606 pp.).
- Chen, J., Kemna, A., Hubbard, S.S., 2008. A comparison between Gauss–Newton and Markov-chain Monte Carlo-based methods for inverting spectral induced-polarization data for Cole–Cole parameters. *Geophysics* 73 (6), F247–F259.
- Cole, K.S., Cole, R.H., 1941. Dispersion and absorption in dielectrics. I Alternating current characteristics *Journal of Chemical Physics* 9, 341–351.
- Dias, C.A., 2000. Developments in a model to describe low-frequency electrical polarization of rocks. *Geophysics* 65, 437–451.
- Fernández-Martínez, J.L., Fernández-Muñiz, Z., Tompkins, M.J., 2012. On the topography of the cost functional in linear and nonlinear inverse problems. *Geophysics* 77, W1–W15. <http://dx.doi.org/10.1190/geo2011-0341>.
- Florsch, N., Llubes, M., Téreygeol, F., Ghorbani, A., Roblet, P., 2011. Quantification of slag heap volumes and masses through the use of induced polarization: application to the Castel-Minier site. *J. Archaeol. Sci.* 38. <http://dx.doi.org/10.1016/j.jas.2010.09.027>.
- Florsch, N., Camerlynck, C., Revil, A., 2012. Direct estimation of the distribution of relaxation times from induced-polarization spectra using a Fourier transform analysis. *Near Surf. Geophys.* 10, 517–531.
- Ghorbani, A., Camerlynck, C., Florsch, N., Cosenza, Ph., Tabbagh, A.I., Revil, A., 2007. Bayesian inference of the Cole–Cole parameters from time and frequency domain induced polarization. *Geophys. Prospect.* 55 (4), 589–605. <http://dx.doi.org/10.1111/j.1365-2478.2007.00627.x>.
- Hansen, P.C., O’Leary, D.P., 1993. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.* 14, 1487–1503.
- Havriliak, S., Negami, A., 1966. A complex plane analysis of α -dispersions in some polymer systems. *J. Polym. Sci.* 14, 99–117. <http://dx.doi.org/10.1002/polc.5070140111>.
- Hilfer, R., 2002. Analytical representations for relaxation functions of glasses. *J. Non-Cryst. Solids* 305, 122–126.
- Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, San Diego.
- Miranda, D.A., Lopez-Rivera, S.A., 2008. Determination of Cole–Cole parameters using only the real part of electrical impedivity measurements. *Physiol. Meas.* 29, 669–683. <http://dx.doi.org/10.1088/0967-3334/29/5/011>.
- Nordsiek, S., Weller, A., 2008. A new approach to fitting induced-polarization spectra. *Geophysics* 73, F235–F245.
- Pelton, W.H., Ward, S.H., Hallof, P.G., Sill, W.R., Nelson, P.H., 1978. Mineral discrimination and removal of inductive coupling with multifrequency IP. *Geophysics* 43, 588–609.
- Pelton, W.H., Sill, W.R., Smith, B.D., 1983. Interpretation of complex resistivity and dielectric data – part I. *Geophys. Trans.* 29, 297–330.
- Revil, A., 2013. On charge accumulation in heterogeneous porous rocks under the influence of an external electric field. *Geophysics* 78, 271–291. <http://dx.doi.org/10.1190/GEO2012-0503.1>.
- Revil, A., Florsch, N., 2010. Determination of permeability from spectral induced polarization data in granular media. *Geophys. J. Int.* 181, 1480–1498. <http://dx.doi.org/10.1111/j.1365-246X.2010.04573.x>.
- Revil, A., Leroy, P., Ghorbani, A., Florsch, N., Niemeijer, A.R., 2006. Compaction of quartz sands by pressure solution using a Cole–Cole distribution of relaxation times. *J. Geophys. Res.* 111, B09205. <http://dx.doi.org/10.1029/2005JB004151>.
- Revil, A., Koch, K., Holliger, K., 2012. Is it the grain size or the characteristic pore size that controls the induced polarization relaxation time of clean sands and sandstones? *Water Resour. Res.* 48, W05602. <http://dx.doi.org/10.1029/2011WR011561>.
- Revil, A., Florsch, N., Camerlynck, C., 2014. Spectral induced polarization porosimetry. *Geophysical Journal International* 198, 1016–1033. <http://dx.doi.org/10.1093/gji/ggu180>.
- Schwarz, G., 1962. A theory of the low-frequency dielectric dispersion of colloidal particles in electrolyte solution. *J. Phys. Chem.* 66, 2636–2642.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM vol. XII ISBN 978-0-898715-72-9, (342 pages).
- Tarantola, A., Valette, B., 1982. Generalized nonlinear inverse problems solved using the least square criterion. *Rev. Geophys. Space Phys.* 20 (2), 219–232.
- Tarasov, A., Titov, K., 2007. Relaxation time distribution from time domain induced polarization measurements. *Geophys. J. Int.* 170, 31–43.
- Telford and Sheriff 1990 (corriger ds le texte, manqué virgule).
- Tikhonov, A.N., Arsenin, V.Y., 1977. *Solutions of Ill-posed Problems*. John Wiley, Washington, D. C.
- Tong, M., Li, L., Wang, W., Jiang, Y., 2006. Determining capillary-pressure curve, pore-size distribution, and permeability from induced polarization of shaley sand. *Geophysics* 71, 33–40.
- Warburg, E., 1899. Über das Verhalten sogenannter unpolarisierbarer Elektroden gegen Wechselstrom (About the behaviour of so-called “impolarizable electrodes” in the present of alternating current). *Ann. Phys. Chem.* 67, 493–499.
- Wong, J., 1979. An electrochemical model of the induced-polarization phenomenon in disseminated sulfide ores. *Geophysics* 44 (7), 1245–1265.