Interval Simulated Annealing Applied to Electrical Impedance Tomography

Thiago de Castro Martins, thiago@usp.br
Marcos de Sales Guerra Tsuzuki, mtsuzuki@usp.br
Computational Geometry Laboratory, Escola Politécnica da USP

Abstract. *Electrical Impedance Tomography (EIT) is a imaging technique that attempts to reconstruct the impedance distribution inside an object from the impedance between electrodes placed at the object's surface. The EIT reconstruction problem can be approached as a non-linear non-convex optimization problem where one tries to maximize the matching between a simulated impedance problem and the observed data. This non-linear optimization problem is often ill-posed, and not very suited to methods that evaluate derivatives of the objective function. It may be approached by Simulated Annealing (SA), but at a large computational cost due to the expensive evaluation process of the objective function, that involves a full simulation of the impedance problem at each iteration. We propose here a variation of SA where the objective function is evaluated only partially, while ensuring boundaries on the behavior of the modified algorithm.*
Keywords: *Simulated Annealing, Electrical Impedance Tomography, Inverse Problem.*

1. Introduction

The non invasive and non destructive imaging of the interior of objects has been a challenge. Electrical impedance tomography (EIT) is one possibility to obtain information from otherwise inaccessible regions within a closed and often opaque volume. EIT estimates the electrical conductivity distribution within the body when a low amplitude current pattern is applied to a body surface and the potential at determined points of that surface is measured through electrodes or, alternatively, when a potential is applied and the current flowing through the surface is measured (Trigo *et al.*, 2004).

The model of the body is based on an elliptic partial differential equation obtained from Maxwell's equations and the electrical conductivity distribution represents the solution of an ill-posed nonlinear inverse problem, meaning that large changes in conductivity at the interior of the body may result in only small potential or current changes at the surface. The main problems with EIT are its sensitivity to electrode positioning, its rather low image resolution due to noise in the measurements and numerical effects of the discretization, and its sensitivity to the boundary geometry.

The two main forms of EIT are dynamic imaging and static imaging yielding differential and absolute images respectively. The images produced by differential imaging represent the conductivity changes of a region between two time intervals (Barber and Brown, 1984). Imaging physiological function within the body largely relies on this technique. This work, is mainly concerned with the reconstruction of static conductivity images which requires more advanced numerical algorithms.

EIT has a wide range of medical applications, detection of acute cerebral stroke (Clay and Frree, 2002), breast cancer (Kao *et al.*, 2006), monitor cardiac activity (Eyüboglu *et al.*, 1989) and monitor lung aeration imposed by mechanical ventilation in critically ill patients (Hua *et al.*, 1993; Trigo *et al.*, 2004). The last application is the main interest of our group, where EIT can be used to implement a protective lung strategy that can result in an improved lung function and survival (Amato *et al.*, 1998).

2. Formulation of the forward problem

The typical forward problem in EIT is given the conductivity distribution $\boldsymbol{\sigma}$ and the current $\boldsymbol{J}$ injected through boundary electrodes, find the potential distribution $\boldsymbol{\phi}$ within Ω and in particular the resulting potentials at the measurement electrodes $\boldsymbol{\phi}_m$. The frequencies used in EIT are low enough so that the quasi-static approximation hold, and thus we can ignore capacitive and inductive effects. Under such quasi-static conditions, the solution of the forward problem is rather simple as it only requires solving the Laplace equation

$$\nabla \left(\sigma \nabla \phi \right) = 0. \tag{1}$$

That is an elliptic equation in the divergence form (Evans, 2010, ch. 6). The boundary conditions are given by the currents J at the domain's frontier

$$\sigma \frac{\partial \phi}{\partial \hat{n}} = J \tag{2}$$

where $\hat{n}$ is the external normal. The convention adopted here is to have positive J values for currents *exiting* the domain. First-derivative boundary conditions such as the formulated in (2) are called *Neumann boundary conditions*. The forward problem with Newmann boundary conditions is particular to scenarios where the boundary currents are known. Another possible scenario is that of known boundary potentials, leading to *Dirichlet boundary conditions*. On boundary conditions,

it is important to notice that integrating (1) over the domain Ω and applying Gauss' theorem, then

$$\oint_{\partial\Omega} \sigma\nabla\phi\cdot\hat{n}\,\mathrm{d}s = 0. \tag{3}$$

Replacing (2) in (3),

$$\oint_{\partial\Omega} J\,\mathrm{d}s = 0$$

that is a necessary condition for (1,2) to have a solution. Physically, it is equivalent to require that the sum of currents entering and leaving the domain is zero.

2.1 Variational Formulation of the forward problem

Consider the functional $F(\phi)$ given by

$$F(\phi) = \frac{1}{2}\int_\Omega \sigma\nabla\phi\cdot\nabla\phi\,\mathrm{d}x - \oint_{\partial\Omega}\phi J\,\mathrm{d}s. \tag{4}$$

If ϕ^* is a solution of (1, 2), then ϕ^* is a minimum of $F(\phi)$. Indeed, let ϕ be an arbitrary potential distribution, and ϕ^* a solution of (1). Then,

$$\int_\Omega (\phi^*-\phi)\left[\nabla(\sigma\nabla\phi^*)\right]\,\mathrm{d}x = 0$$

integrating by parts,

$$\int_\Omega \sigma\nabla(\phi^*-\phi)\cdot\nabla\phi^*\,\mathrm{d}x - \oint_{\partial\Omega}(\phi^*-\phi)\,\sigma\nabla\phi^*\cdot\hat{n}\,\mathrm{d}s = 0$$

replacing (2) on the second term and expanding,

$$\int_\Omega \sigma\nabla\phi^*\cdot\nabla\phi^*\,\mathrm{d}x = \int_\Omega \sigma\nabla\phi^*\cdot\nabla\phi\,\mathrm{d}x - \oint_{\partial\Omega} J(\phi-\phi^*)\,\mathrm{d}s = 0.$$

Applying the Cauchy inequality [1] to the first term and collecting on the left side the terms in ϕ^*, we have

$$\frac{1}{2}\int_\Omega \sigma\nabla\phi^*\cdot\nabla\phi^*\,\mathrm{d}x - \oint_{\partial\Omega} J\phi^*\,\mathrm{d}s \le \frac{1}{2}\int_\Omega \sigma\nabla\phi\cdot\nabla\phi\,\mathrm{d}x - \oint_{\partial\Omega} J\phi\,\mathrm{d}s$$

for any ϕ, proving that ϕ^* is a minimum of (4). Applying the *Generalized Dirichlet Principle* (Evans, 2010, p. 456-467), one can show that the reverse is also true, that is, if a potential distribution ϕ^* is a minimum of (4), then it is also a solution of (1, 2). In general lines, if ϕ^* is a minimum of F, then it's variation ∂F is zero at ϕ^*

$$\partial F(\phi)|_{\phi=\phi^*} = \partial\left(\frac{1}{2}\int_\Omega \sigma\nabla\phi\cdot\nabla\phi\,\mathrm{d}x - \oint_{\partial\Omega}\phi J\,\mathrm{d}s\right)\Bigg|_{\phi=\phi^*} = \int_\Omega \nabla(\partial\phi)\,\sigma\nabla\phi^*\,\mathrm{d}x - \oint_{\partial\Omega}\partial\phi J\,\mathrm{d}s = 0.$$

Where $\partial\phi$ is an infinitesimal arbitrary variation of the function ϕ^*. Applying Green's formulas to the first term,

$$0 = \int_\Omega \nabla\left[\partial\phi\,(\sigma\nabla\phi^*)\right]\,\mathrm{d}x - \int_\Omega \partial\phi\nabla(\sigma\nabla\phi^*)\,\mathrm{d}x - \oint_{\partial\Omega}\partial\phi J\,\mathrm{d}s.$$

Applying Gauss's theorem to the first term,

$$\int_\Omega \partial\phi\nabla(\sigma\nabla\phi^*)\,\mathrm{d}x = \oint_{\partial\Omega}\partial\phi\,(\nabla\phi^*\cdot\hat{n} - J)\,\mathrm{d}s \tag{5}$$

This equation holds true for *any* $\partial\phi$. In particular, if we take $\partial\phi = 0$ in $\partial\Omega$,

$$\int_\Omega \partial\phi\nabla(\sigma\nabla\phi^*)\,\mathrm{d}x = 0 \Longrightarrow \nabla(\sigma\nabla\phi^*) = 0$$

showing that ϕ^* satisfies (1). On the other hand, if $\partial\phi \neq 0$ in $\partial\Omega$ then, replacing (1) in (5),

$$\oint_{\partial\Omega}\partial\phi\,(\nabla\phi^*\cdot\hat{n} - J)\,\mathrm{d}s = 0 \Longrightarrow \sigma\frac{\partial\phi}{\partial\hat{n}} = J$$

showing that ϕ^* satisfies (2). This is somehow surprising, as (4) does not even require ϕ to have a second order derivative (it only requires for the two integrands to be integrable). The formulation of a differential equation as a functional optimization problem is known as the *weak* or *variational* formulation.

[1] $\vec{u}\cdot\vec{v} \le \frac{1}{2}(\vec{u}\cdot\vec{u} + \vec{v}\cdot\vec{v})$

2.2 Finite Element Method applied to the forward problem

Analytical solutions for (1, 2) for arbitrary domains Ω are not known. The Finite Element Method (FEM) is a process to obtain approximate numerical solutions for Partial Differential Equations (PDEs) such as those studied in this work (Castro, 2006; Bathe, 1995). It can be derived from the variational form of the problem described in section 2.1. The basic idea is to adopt as approximated solution of (1,2) a minimum of (4) when the potential distributions ϕ are restricted to a linear subspace generated by a base of functions. As such, let $\mathbf{\Psi} = [\psi_1, \psi_2, \ldots, \psi_n]^T$ a finite vector of n linearly independent functions such that $\psi_i \in C\left(\mathbb{R}^2\right)$. The domain $\mathcal{A}$ of functions considered by FEM is given by

$$\mathcal{A} = \left\{ u : \mathbb{R}^2 \to \mathbb{R} \mid u = \mathbf{x}^T \cdot \mathbf{\Psi}, \mathbf{x} \in \mathbb{R}^n \right\}. \tag{6}$$

An approximated solution $\tilde{\phi}$ can be written as

$$\tilde{\phi} = \mathbf{\Phi}^T \cdot \mathbf{\Psi} \tag{7}$$

where $\mathbf{\Phi} = [\phi_1, \phi_2, \ldots, \phi_n]^T$ is the coefficient vector that defines $\tilde{\phi}$. As stated above, $\tilde{\phi}$ is the function in $\mathcal{A}$ that minimizes (4). Replacing (7) in (4) we have

$$F\left(\tilde{\phi}\right) = \frac{1}{2}\int_\Omega \sigma \left(\mathbf{\Phi}^T \cdot \nabla\mathbf{\Psi}\right) \cdot \left(\mathbf{\Phi}^T \cdot \nabla\mathbf{\Psi}\right) \, \mathrm{d}x - \oint_{\partial\Omega} \mathbf{\Phi}^T \cdot \mathbf{\Psi} \cdot \boldsymbol{J} \, \mathrm{d}s.$$

Using $\otimes$ as notation for the outer product [2] and grouping the terms in $\nabla\mathbf{\Psi}$,

$$F\left(\tilde{\phi}\right) = \frac{1}{2}\int_\Omega \sigma \, \mathbf{\Phi}^T \cdot (\nabla\mathbf{\Psi} \otimes \nabla\mathbf{\Psi}) \cdot \mathbf{\Phi} \, \mathrm{d}x - \oint_{\partial\Omega} \mathbf{\Phi}^T \cdot \mathbf{\Psi} \cdot \boldsymbol{J} \, \mathrm{d}s.$$

As the vector $\mathbf{\Phi}$ is independent of x,

$$F\left(\tilde{\phi}\right) = \frac{1}{2}\mathbf{\Phi}^T \cdot \overbrace{\left(\int_\Omega \sigma \left(\nabla\mathbf{\Psi} \otimes \nabla\mathbf{\Psi}\right) \, \mathrm{d}x\right)}^{n\times n} \cdot \mathbf{\Phi}^T - \overbrace{\left(\oint_{\partial\Omega} \mathbf{\Psi} \cdot \boldsymbol{J} \, \mathrm{d}s\right)^T}^{n\times 1} \cdot \mathbf{\Phi}.$$

Notice that the integral in the first term is now a $n \times n$ array, and that in the second term is a vector of n components. Defining the *Stiffness Matrix*[3] of the system $\mathbf{K}(\sigma)$ as

$$\boldsymbol{K}(\sigma) = \int_\Omega \sigma \left(\nabla\mathbf{\Psi} \otimes \nabla\mathbf{\Psi}\right) \, \mathrm{d}x$$

and the *Load Vector* $\boldsymbol{C}$

$$\boldsymbol{C} = \oint_{\partial\Omega} \mathbf{\Psi} \cdot \boldsymbol{J} \, \mathrm{d}s$$

we have

$$F\left(\tilde{\phi}\right) = \frac{1}{2}\mathbf{\Phi}^T \cdot \boldsymbol{K}(\sigma) \cdot \mathbf{\Phi} - \boldsymbol{C}^T \cdot \mathbf{\Phi}. \tag{8}$$

By definition. $K(\sigma)_{ij} = \int_\Omega \sigma \left(\nabla\psi_i \otimes \nabla\psi_j\right) \, \mathrm{d}x$ is symmetric and positive definite (it suffices to notice that $\mathbf{\Phi}^T \cdot \boldsymbol{K}(\sigma) \cdot \mathbf{\Phi}^T$ is equivalent to the first term of (4), that is necessarily positive). As such, it is easy to show that (8) has a minimum (see for instance Shewchuk (1994)) given by the solution of

$$\boldsymbol{K}(\sigma) \cdot \mathbf{\Phi} - \boldsymbol{C} = 0 \tag{9}$$

that is a linear system that can be efficiently solved by numerical methods.

[2] The outer product or tensor product between two vectors $u_{1\times m}$ e $v_{1\times m}$ is given by $u \otimes v = P_{n\times m}$ such that $P_{ij} = u_i v_j$. It can be easily shown that $\left(a^T b\right)\left(c^T d\right) = a^T \left(b \otimes c\right) d$

[3] The denominations *Stiffness Matrix* and *Load Vector* come from the traditional application of FEM to the Field of Structural problems

2.3 The Inverse Problem as an Optimization Problem

The EIT inverse problem is the problem of, given the electric currents and tensions on the boundary of the domain for multiple patterns, find the distribution of electric conductivity in the domain interior. Since there are known methods for efficiently solve the forward problem (such as FEM), one possible approach to the inverse problem is to look at it as an optimization problem, where the optimization variables are a parametrization of the conductivity inside the domain (for instance, the domain is divided in continuous segments and the conductivity inside each segment — presumed to be constant — becomes a parameter) and the optimization function is some measure of how the solution of the forward problem applied to the conductivity distribution produced by the optimization variables matches the measured data. One possible objective function E is the Euclidean distance between the measured electric potentials ϕ_m^i and the calculated potentials ϕ_c^i for all the applied current patterns

$$E = \sqrt{\Sigma |\phi_m^i - \phi_c^i|^2}. \tag{10}$$

An example of such approach is in (Mello *et al.*, 2008), where the objective function 10 is minimized by Sequential Linear Programming yielding estimations of the conductivity distribution. It was pointed by Mello *et al.* (2008) that the solution of this optimization problem by methods that require the computation of the objective function gradients is difficult since – as the problem is intrinsically ill–conditioned – the errors on such computations can become quite large.

That is why the interest on Simulated Annealing (SA) applied to EIT is increasing, as it requires no evaluation of objective function derivatives (in fact, as we will show, it does not even require a complete computation of the objective function).

3. Simulated Annealing

SA is a hill-climbing local exploration optimization heuristic, which means it can skip local minima by allowing the exploration of the space in directions that lead to a local increase on the cost function. It sequentially applies random modifications on the evaluation point of the cost function. If a modification yields a point of smaller cost, it is automatically kept. Otherwise, the modification also can be kept with a probability obtained from the Boltzman distribution

$$P(\Delta E) = e^{-\frac{\Delta E}{kT}} \tag{11}$$

where $P(\Delta E)$ is the probability of the optimization process to keep a modification that incurs an increase ΔE of the cost function. k is a parameter of the process (analogous to the Stefan–Boltzman constant) and T is the instantaneous "temperature" of the process. This temperature is defined by a cooling schedule, and it is the main control parameter of the process. The probability of a given state decreases with its energy, but as the temperature rises, this decrease (the slope of the curve $P(\Delta E)$) diminishes.

4. Applying Simulated Annealing to EIT

As seen in section 2.3, the EIT inverse problem can be formulated as an optimization problem, and as such, can be approached with SA. Herrera *et al.* (2007) minimized objective function (10) calculated in a model and measured on real subjects with SA and by doing so, managed to reconstruct very accurate conductivity distributions of the body, but at a very high computational cost. This is unsurprising, as each step of the SA involves the solution of a full FEM problem in order to evaluate the objective function.

4.1 Incomplete Evaluation of the Objective Function

As the evaluation of the objective function is responsible for the bulk of the SA computational cost, it is interesting to look for means to reduce its cost. In particular, we will study how the SA process behaves in the presence of a partial evaluation of the objective function. We will use the objective function given by (10). The calculated potentials ϕ_c^i for each current pattern are obtained from a FEM algorithm whose kernel is a solver for the linear system given by (9). Notice that there are multiple current patterns C^i, leading to multiple calculated potentials Φ^i. Let us suppose that there is a way to iteratively solve the linear system such that at each iteration we have an estimative of the solution and a (decreasing) upper limit on the l_2 norm of the difference between the estimative and the correct solution. Then, if $d^i = \left|\phi_m^i - \phi_c^i\right|$ is the component of the objective function for a given current pattern, $\tilde{d}^i = \left|\phi_m^i - \tilde{\phi}_c^i\right|$ is its estimate obtained with the partial solution $\tilde{\phi}_c$ and $\epsilon^i \geq \left|\phi_c^i - \tilde{\phi}_c^i\right|$ is an upper boundary on the error of the solver algorithm, we have

$$\begin{aligned} d^i \leq d_{max}^i &= \tilde{d}^i + \epsilon^i \\ d^i \geq d_{min}^i &= \max\left(0, \tilde{d}^i - \epsilon^i\right) \end{aligned}$$

and of course,

$$\begin{aligned} \tilde{E} &= \sqrt{\Sigma\left(\tilde{d}^i\right)^2} \\ E \leq E_{max} &= \sqrt{\Sigma(d^i_{max})^2} \\ E \geq E_{min} &= \sqrt{\Sigma\left(d^i_{min}\right)^2}. \end{aligned}$$

The upper and lower boundaries for E converge to the exact value as the number of iterations of the multiple solver algorithms increases. As SA is sensible only to relative variations of the objective function those boundaries must be converted to boundaries of the variation of E. If $\Delta E = E^j - E^i$ is the variation of the objective function when the SA process moves from one solution x^j to another x^{j+1} then, by interval arithmetic,

$$\begin{aligned} \Delta\tilde{E} &= \tilde{E}^j - \tilde{E}^{j+1} \\ \Delta E \leq \Delta E_{max} &= E^j_{max} - E^{j+1}_{min} \\ \Delta E \geq \Delta E_{min} &= E^j_{min} - E^{j+1}_{max}. \end{aligned} \tag{12}$$

Again, those boundaries for ΔE converge to the exact value, but this time, it is necessary to increase the number of iterations for both the evaluation of x^j and x^{j+1}.

From that it can be seen that a SA process that would use the partial solutions $\tilde{\phi}_c$ for its linear systems instead of the exact ones ϕ_c would have a limited probability of diverging from the process that uses the exact solutions. Indeed, by imposing P_{err} as an upper limit on the probability of the process taking a "wrong" decision (rejecting a solution when it should accept it or accepting when it should reject), we can create conditions for the boundaries given in (12)

$$e^{-\Delta E_{max}/kt} \geq \begin{cases} 1 - P_{err} & \text{if } \Delta\tilde{E} \leq 0, \\ e^{-\Delta\tilde{E}/kt} - P_{err} & \text{if } \Delta\tilde{E} > 0 \end{cases} \tag{13}$$

$$e^{-\Delta E_{min}/kt} \leq min(1, P_{err} + e^{-\Delta\tilde{E}/kt}). \tag{14}$$

The conditions given by (13) and (14) do not translate directly into stopping criteria for the solver algorithm, as ΔE_{max} and ΔE_{min} are the partial solutions of two separate sets of FEM problems, and in general, they are not reachable by tightening the boundaries in just one set. As such, every solution must have its objective function evaluation process stored so it may be possible to continue it from where it has stopped.

In fact, it is often computationally more efficient to improve the boundaries on the evaluation of the *previous* SA solution than the current one. Since CG has an asymptotically geometric convergence behavior and presuming that both the previous SA solution and the current one have similar convergence rates (a reasonable assumption is that the **K** matrix varies very little from one iteration to another and a preconditioner equalizes even more the convergence rate), a simple heuristic for picking which solution to evaluate further is to just pick the one with larger boundaries.

5. Conjugated Gradients Algorithm with estimates on the l_2 norm of the error

The CG Algorithm is an iterative algorithm for solving systems of the form

$$\mathbf{A} \cdot \mathbf{x} - \mathbf{b} = 0 \tag{15}$$

when **A** is a symmetric positive definite matrix. It is well suited for solving linear systems like (9). Indeed, a way to derive the algorithm is from the minimization of quadratic forms like (8) (see for instance (Golub and Van Loan, 199, sec. 10.2) or (Shewchuk, 1994)). For this work, it is more interesting to consider its derivation from the Lanczos Algorithm. The Lanczos Algorithm is a decomposition algorithm for symmetric positive definite matrices in the form

$$\mathbf{A} = \mathbf{V} \cdot \mathbf{T} \cdot \mathbf{V}^T \tag{16}$$

where **V** is an orthonormal matrix and **T** is a tridiagonal symmetric matrix. In (Meurant, 2006) it is shown how **V** is built as an orthonormal basis for the Krylov subspace

$$\mathcal{K}(\mathbf{A}, \mathbf{v}) = \text{span}\left\{\mathbf{v}, \mathbf{A} \cdot \mathbf{v}, \mathbf{A}^2 \cdot \mathbf{v}, \ldots, \mathbf{A}^{n-1} \cdot \mathbf{v}\right\}.$$

Indeed, applying the Grand-Schmidt orthogonalization process to $\mathcal{K}(\mathbf{A}, \mathbf{v})$, we have:

$$\begin{aligned}
\mathbf{v}_1 &= {}^{\mathbf{v}}/_{\|\mathbf{v}\|} \\
t_{i,j} &= \mathbf{v}_j^T \cdot \mathbf{A} \cdot \mathbf{v}_i,\ j \leq k \\
\bar{\mathbf{v}}_i &= \mathbf{A} \cdot \mathbf{v}_i - \sum_{j=1}^{n} t_{j,i}\mathbf{v}_j \\
t_{i+1,i} &= \|\bar{\mathbf{v}}_i\| \\
\mathbf{v}_{i+1} &= {}^{\bar{\mathbf{v}}_i}/_{t_{i+1,i}}.
\end{aligned}$$

Let $\mathbf{V}_k$ be a $n \times k$ matrix collecting the $\mathbf{v}_i$ vectors generated in the first k iterations and $\mathbf{T}_k$ a $k \times k$ matrix whose coefficients are $t_{i,j}$ (notice that every coefficient below the first subdiagonal is zero). It can be shown that $\mathbf{T}_k$ is symmetric and consequentially, tridiagonal. Then the decomposition in (16) can be obtained taking $\mathbf{V} = \mathbf{V}_n$ and $\mathbf{T} = \mathbf{T}_n$.

The Lanczos decomposition can be used to solve linear systems like (15). Indeed, let $\mathbf{x}_0$ be an initial guess and $\mathbf{r}_0 = \mathbf{A} \cdot \mathbf{x}_0 - \mathbf{b}$ the initial residue. Then, if $\mathbf{r}_0$ is taken as the initial vector for the Krylov subspace $\mathcal{K}(\mathbf{A}, \mathbf{r}_0)$. An iterative solution may be computed using the vectors in $\mathbf{V}_k$. The solution at iteration k is given by (Golub and Meurant, 2010)

$$\mathbf{x}_k = \mathbf{x}_0 + \mathbf{V}_k \cdot \mathbf{y}_k \tag{17}$$

the vector $\mathbf{y}_k$ is taken so the resulting residual $\mathbf{r}_k = \mathbf{A} \cdot \mathbf{x}_k - \mathbf{b}$ is orthogonal to $\mathbf{V}_k$. It is clear that if such values of $\mathbf{y}_k$ can be found, then $\mathbf{r}_n = 0$ and $\mathbf{x}_n$ is a solution of (15). By imposing $\mathbf{V}_k = 0$ we have the condition $\mathbf{T}_k \cdot \mathbf{y}_k = \mathbf{V}_k^T \cdot \mathbf{r}_0$. Since the vector $\mathbf{r}_0$ is taken as the base of the Krylov subspace, $\mathbf{V}_k^T \cdot \mathbf{r}_0 = \|\mathbf{r}_0\| \mathbf{e}^1$, where $\mathbf{e}^1$ is the first column of the identity matrix of rank n. The vector $\mathbf{y}_k$ is then given by $\mathbf{y}_k = \|\mathbf{r}_0\| \mathbf{T}_k^{-1} \cdot \mathbf{e}^1$, leading to

$$\mathbf{x}_k = \mathbf{x}_0 + \|\mathbf{r}_0\| \mathbf{V}_k \cdot \mathbf{T}_k^{-1} \cdot \mathbf{e}^1.$$

$\mathbf{T}_k$ in turn can be decomposed as $\mathbf{L}_k \cdot \mathbf{D}_k^{-1} \cdot \mathbf{L}_k^T$, where $\mathbf{D}_k$ is a diagonal matrix and $\mathbf{L}_k$ is a lower bidiagonal matrix whose main diagonal is the same as $\mathbf{D}_k$ and its subdiagonal is equal to the subdiagonal of $\mathbf{T}_k$. Through this decomposition it can be shown (Golub and Meurant, 2010, pg 46-49) that $\mathbf{x}_k$ can be computed iteratively with only a single matrix $\times$ vector product per iteration. The full algorithm follows:

Algorithm 1 Solve $\mathbf{A} \cdot \mathbf{x} - \mathbf{b} = 0$

Input: $\mathbf{A}$ a $n \times n$ symmetric positive definite matrix, $\mathbf{b}$ and $\mathbf{x}_0$
Output: $\mathbf{x}$ such that $\mathbf{A} \cdot \mathbf{x} - \mathbf{b} = 0$

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A} \cdot \mathbf{x}_0$
$\mathbf{p}_0 = \mathbf{r}_0$
$\gamma_0 = {}^{\mathbf{r}_0^T \cdot \mathbf{r}_0}/_{\mathbf{r}_0^T \cdot \mathbf{A} \cdot \mathbf{r}_0}$
for $i = i$ to n **do**
  $\mathbf{x}_i = \mathbf{x}_{i-1} + \gamma_{i-1}\mathbf{p}_{i-1}$
  $\mathbf{r}_i = \mathbf{r}_{i-1} - \gamma_{i-1}\mathbf{A} \cdot \mathbf{p}_{i-1}$
  $\beta_i = {}^{\mathbf{r}_i^T \cdot \mathbf{r}_i}/_{\mathbf{r}_{i-1}^T \cdot \mathbf{r}_{i-1}}$
  $\mathbf{p}_i = \mathbf{r}_i + \beta_i \mathbf{p}_{i-1}$
  $\gamma_i = {}^{\mathbf{r}_i^T \cdot \mathbf{r}_i}/_{\mathbf{p}_i^T \cdot \mathbf{A} \cdot \mathbf{p}_i}$
end for
output $\mathbf{x}_n$

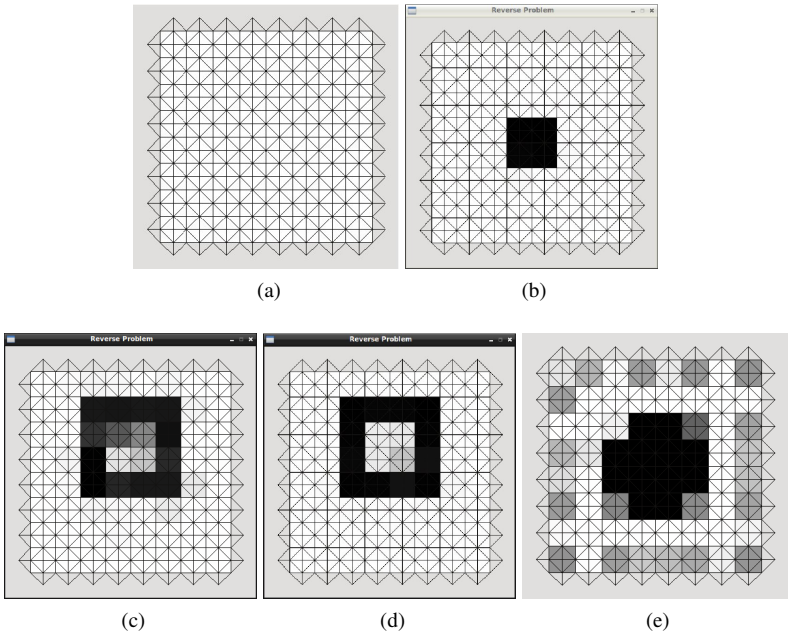
5.1 The l_2 error estimate

Let $\boldsymbol{\epsilon}_k = \mathbf{x}_k - \mathbf{A}^{-1} \cdot \mathbf{b}$ be the error of the algorithm at iteration k. Noticing that $\boldsymbol{\epsilon}_k = \mathbf{x}_k - \mathbf{x}_0 - \mathbf{A}^{-1} \cdot \mathbf{r}_0$ and applying the relation (17),

$$\|e_k\|^2 = \mathbf{r}_0^T \cdot \mathbf{A}^{-2} \cdot \mathbf{r}_0 - 2\mathbf{r}_0^t \cdot \mathbf{A}^{-1} \cdot \mathbf{V}_k \cdot \mathbf{y}_k + \mathbf{y}_k^T \cdot \mathbf{V}_k \cdot \mathbf{y}_k.$$

By applying the decomposition of $\mathbf{T}_k$ (Golub and Meurant, 2010, pg 69),

$$\|e_k\|^2 = \|\mathbf{r}_0\|^2 \left(\left(\mathbf{T}_n^{-2}\right)_{1,1} - \left(\mathbf{T}_k^{-2}\right)_{1,1} \right) - 2 \|\boldsymbol{\epsilon}_k\|_A^2 \frac{\left(\mathbf{T}_k^{-2}\right)_{k,1}}{\left(\mathbf{T}_k^{-1}\right)_{k,1}}$$



(a) (b) (c) (d) (e)

Figure 1: (a) Simulation Domain. (b) Final result for the "Central Dot" problem. (Higher conductance is brighter). (c) Final result for the "Box" problem. (d) Final result for the "Box" problem with modified neighborhood heuristic. (e) Evaluating the effect of FEM mesh errors.

where $\|\boldsymbol{\epsilon}_k\|_A^2 = \boldsymbol{\epsilon}_k^A \boldsymbol{\epsilon}_k$ is the A-norm of the error at iteration k. Since $\mathbf{T}_k$ is positive definite, the last term will be always negative, and will be disregarded here, as we are interested in upper bounds for the norm of the error. Of course, the $\mathbf{T}_n$ matrix is unknown until the algorithm has reached the n-th iteration. Considering that the convergence of the CG is geometric, by introducing a delay d such that $\|\boldsymbol{\epsilon}_{k-d}\|^2 \gg \|\boldsymbol{\epsilon}_k\|^2$, we may estimate an upper bound for $\|\boldsymbol{\epsilon}_{k-d}\|_A^2$ at iteration k as

$$\|\mathbf{r}_0\|^2 \left(\left(\mathbf{T}_k^{-2}\right)_{1,1} - \left(\mathbf{T}_{k-d}^{-2}\right)_{1,1} \right).$$

Notice the value will be obtained d iterations "late", but as the norm of the error in the CG is monotonically decreasing, the error estimate for a few iterations earlier is an upper boundary for the error in the current iteration. The value $\left(\mathbf{T}_k^{-2}\right)_{1,1}$ can be iteratively computed in parallel with the CG using a QR factorization of the matrix $\mathbf{T}_k$ (see Meurant (2005) for details, or Ortega and Kaiser (1963) for more general results on the QR decomposition of tridiagonal matrices).

6. Results

To evaluate the viability of SA with partial evaluation applied to the EIT problem, a simple implementation was built using a simulated domain. Our simulated domain shown in Fig. 1(a) is a square, measuring 16×16 units, composed of 76 triangular elements with 320 nodes, of whose 32 are electrodes. The electric conductance domain was divided in 64 2×2 squares (notice the potential discretization and the conductance discretization is not the same). The conductance in each element is defined in the interval $[1, 2]$. The electrodes are modeled as 2×1 rectangles, the larger edge being in contact with the domain. The opposite edge is modeled as being in contact with an infinite conductivity element and consequently the potential across it is constant (the triangular shapes seen in Fig. 1 are thus misleading, they just reflect the fact that the outer edge is collapsed into a single potential node in the FEM model). The first electrode (the topmost electrode in the left column) was defined as the ground. We applied 31 simulated current patterns, corresponding to an unitary current entering each electrode – but the ground – and leaving through the other ground node. The FEM problems were solved using a preconditioned CG method, with Incomplete Choleski decomposition preconditioning (Kershaw, 1978). As suggested by (Meurant, 2005), the pre–conditioner l_∞ norm was used for estimation of the error l_2 norm. The 64 conductivity parameters are generated by SA. The neighborhood heuristic used was taken from (Martins and Tsuzuki, 2009), changing only a single conductivity parameter at each iteration and reducing the modifications on parameters that lead to rejected solutions. The divergence probability P_{err} was arbitrarily defined as $1/100$.

The first evaluated problem was a simple domain with conductivity 2 in all its extension but the 4 central squares, which have conductivity 1. Fig. 1b shows the final result of the SA process. It can be seen that, despite the use of partial evaluations of the objective function, it still converges to the correct solution.

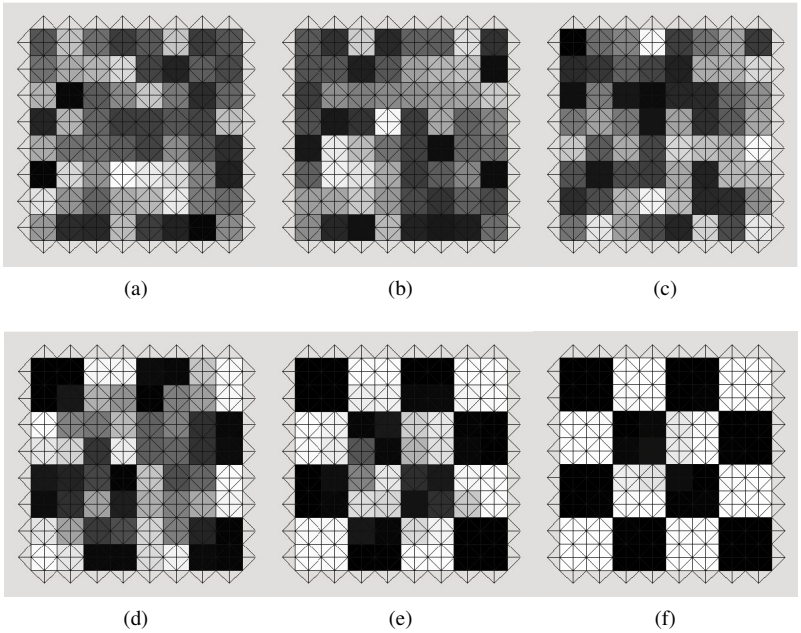


Figure 2: Convergence behavior for the "Checkerboard" problem.

This is a non–convex conductance distribution, with a high impedance (conductivity equal 1) 8×8 box containing a 4×4 low impedance box (conductivity equal 2). As the electric current naturally avoid the box interior, this is a highly ill–conditioned problem, posing a challenge to an partial evaluation approach as ours. Fig. 1c shows the final result of the optimization process. As it can be seen, while the external shape of the high–impedance region was correctly determined, the process got stuck on a local minimum while determining the conductivity of the box interior. We suspected that our initial neighborhood heuristic was to blame, and adopted a less restrictive heuristic where, in addition to modifications to single conductivity values, a "conductivity transport" between two adjacent cells was allowed. The amount of conductivity to be transferred from one cell to another also followed the adaptive heuristic proposed by (Martins and Tsuzuki, 2009). The result with the new neighborhood heuristic is shown in Fig. 1d. As it can be seen, the initial blame on the neighborhood heuristic seems justified, as the conductivity on the interior of the box is accurately determined.

As our goal is to estimate the impact of a partial objective function evaluation on the SA, for all the above problems the observations of current patterns were simulated using the same mesh used by the SA to recreate the impedance (thus compensating FEM errors). It is still interesting to see how the process behaves in the presence of errors introduced by the FEM discretization. For such, a new simulated domain was created using a new mesh with twice the resolution of the old one, composed of 2432 triangular elements. A quick simulation shows that the FEM errors can be quite significant. Estimating the electric impedance between two adjacent electrodes in the middle of one side of the domain (with conductivity 1 in all its extension) using the coarse mesh yields a result of 1.42, while the same simulation using the fine mesh yields a result of 1.56, suggesting that the coarse mesh tends to underestimate the impedance of the electrodes by almost 10%. A problem with a 8×8 square high impedance region was simulated using the fine mesh then the coarse mesh was used by the SA to attempt to reconstruct the conductivity distribution. The results are shown in Fig. 1e. While the image was reasonably reconstructed, there are – as predicted – impedance artifacts in front of the electrodes (caused by the underestimation of electrode impedance by the coarse mesh). Also there is visible noise across the border of the high-impedance region, particularly on its corners.

6.1 Checkerboard Problem and a qualitative observation of the convergence

To retrieve more information from the convergence of the process, we proposed a complex problem – a 4×4 "Checkerboard" conductivity pattern, alternating high conductivity (2) and low conductivity (1). The SA process has accurately reconstructed the conductivity distribution. The examination of intermediary results, displayed on Fig. 2 and **??** shows that the impedance distribution is reconstructed from the outside towards the interior of the domain. While further studies are required, it is reasonable to presume that the objective function is more sensible to modifications on external cells.

Those results suggest the partial evaluation objective function can be as effective as the full evaluation for the SA, but how efficient it is? The figure **??** shows the average number of CG iterations used by the process at each temperature.

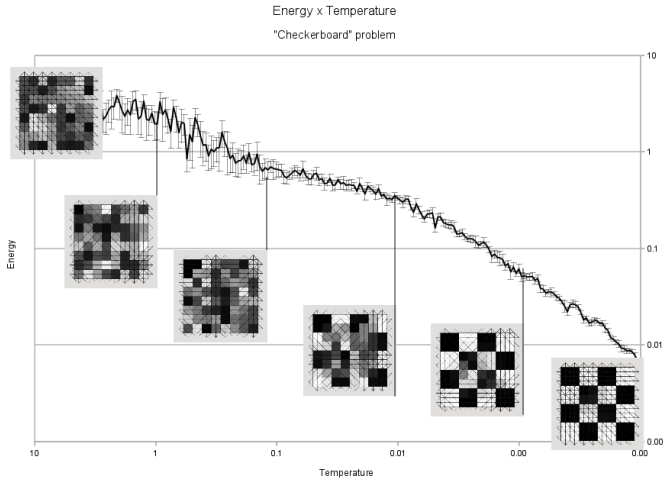


Figure 3: Energy (Objective Function) for the "Checkerboard" problem (error bars represent the standard deviation at a fixed temperature).

Considering that the system has 320 nodes, it is remarkable that the system is able to achieve those results while using in average less than 20 iterations of CG. More interesting is the evolution of the iteration number as the optimization progresses. At high temperatures, the high kt and $\Delta\tilde{E}$ values lead to relatively loose conditions in (13) and (14), reachable by few iterations of the CG algorithm. As the temperature diminishes, lower kt and $\Delta\tilde{E}$ values (as the evaluated solutions become closer) impose through (13) and (14) tighter conditions, leading to a higher number of CG iterations. It may be then a bit surprising that as the process converges towards the final solution the number of CG iterations reduce again. This is explained by two factors.

First, the "initial guess" for the CG algorithm is obtained from the solution obtained at the previous solution of the SA algorithm. Second, the adaptive neighborhood heuristic reduces dramatically the modifications on the conductivity distribution as the process converges towards a global optimum (Martins and Tsuzuki, 2009). Combining those two factors, it can be seen that in the final convergence of the optimization process, the previous FEM solutions are initial guesses good enough to compensate for the increasingly severer conditions on the boundaries of ΔE.

7. Conclusions

We proposed here a new approach to the application of SA to the EIT inverse problem, where the optimization process is allowed to proceed with just a partial evaluation of the objective function. We showed how in the particular case of EIT it is possible to enforce upper bounds on the probability of the modified SA process to deviate from a regular one (with full evaluation of the objective function) by converting those bounds in stopping criteria for the inner CG algorithms used to solve the forward EIT problem. Initial results show that while the partial evaluation of objective functions do not compromise the convergence of the SA algorithm, it has a great potential of improving its efficiency.

8. ACKNOWLEDGEMENTS

Thiago de Castro Martins was supported by FAPESP (Grant 2009/14699–0) and Marcos de Sales Guerra Tsuzuki was partially supported by the CNPq (Grants 304.258/2007–5 and 309570/2010–7). This research was supported by FAPESP (Grant 2009/07173–2).

9. REFERENCES

Amato, M.B.P., Barbas, C.S.V., Medeiros, D.M., Magaldi, R.B., Schettino, G.P., Lorenzi-Filho, G., Kairalla, R.A., Deheinzelin, D., Munoz, C., Oliveira, R., Takagaki, T.Y. and Carvalho, C.R.R., 1998. "Effects of a protective ventilation strategy on mortality in the acute respiratory distress syndrome". *New Engl J Med*, Vol. 338, pp. 347–354.

Barber, D.C. and Brown, B.H., 1984. "Applied potencial tomography". *J Phys E Sci Instrum*, Vol. 17, pp. 723–733.

Bathe, K.J., 1995. *Finite Element Procedures (Part 1-2)*. Prentice Hall.

Castro, Sobrinho, A.d.S., 2006. *Introdução ao Método dos Elementos Finitos*. Ltda., Editora Ciência Moderna, Rio de

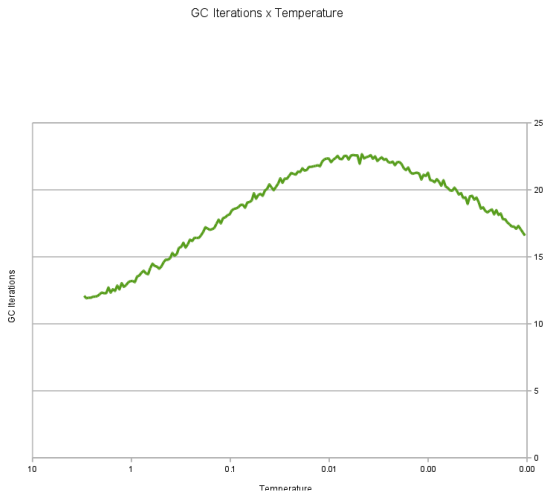


Figure 4: CG Iterations x Temperature for the "Checkerboard" problem.

Janeiro. In Portuguese.

Clay, M.T. and Frree, T.C., 2002. "Weighted regularization in electrical impedance tomography with applications to accute cerebral stroke". *IEEE T Med Imag*, Vol. 21, pp. 629–637.

Evans, L.C., 2010. *Partial Differential Equations: Second Edition (Graduate Studies in Mathematics)*, Vol. 19 of *Graduate Series in Mathematics*. American Mathematical Society, Province, Rhode Island, 2nd edition.

Eyüboglu, B.M., Brown, B.H. and Barber, D.C., 1989. "In vivo imaging of cardiac related impedance changes". *IEEE Eng Med Biol*, Vol. 8, pp. 39–45.

Golub, G.H. and Meurant, G., 2010. *Matrices, Moments and Quadrature with Applications*. Princeton University Press, Princeton, NJ, USA.

Golub, G.H. and Van Loan, C.F., 199. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, 3rd edition.

Herrera, C.N.L., Vallejo, M.F.M., Moura, F.S., Aya, J.C.C. and Lima, R.G., 2007. "Electrical impedance tomography algorithm using simulated annealing search method". In *Proc COBEM*. ABCM, Brasília.

Hua, P., Woo, E.J., Webster, J.G. and Tompkins, W.J., 1993. "Finite element modeling of electrode-sking contact impedance in electrical impedance tomography". *IEEE Eng Med Biol*, Vol. 40, pp. 335–343.

Kao, T.J., Isaacson, D., Newell, J.C. and Saulnier, G.J., 2006. "A 3d reconstruction algorithm for eit using a handheld probe for breast cancer detection". *Physiol Meas*, Vol. 27, pp. S1–S11.

Kershaw, S., 1978. "The incomplete Cholesky-Conjugate Solution of Systems". *J Comput Phys*, Vol. 65, pp. 43–65.

Martins, T.C. and Tsuzuki, M.S.G., 2009. "Placement over containers with fixed dimensions solved with adaptive neighborhood simulated annealing". *B Pol Acad Sci Techn Sci*, Vol. 57, pp. 273–280.

Mello, L.A.M., de Lima, C.R., Amato, M.B.P., Lima, R.G. and Silva, E.C.N., 2008. "Three-dimensional electrical impedance tomography: a topology optimization approach." *IEEE T Bio-Med Eng*, Vol. 55, No. 2 Pt 1, pp. 531–40.

Meurant, G., 2005. "Estimates of the l2 norm of the error in the conjugate gradient algorithm". *Numerical Algorithms*, Vol. 40, No. 2, pp. 157–169. ISSN 1017-1398.

Meurant, G., 2006. *The Lanczos and Conjugate Gradient Algorithms: From Theory to Finite Precision Computations*. Society for Industrial and Applied Mathematics.

Ortega, J.M. and Kaiser, H.F., 1963. "The LLT and QR methods for symmetric tridiagonal matrices". *The Computer Journal*.

Shewchuk, J.R., 1994. "An Introduction to the Conjugate Gradient Method Without the Agonizing Pain". Technical report, Carnegie Mellon University, Pittsburgh, PA, USA.

Trigo, F.C., Gonzales-Lima, R. and Amato, M.B.P., 2004. "Electrical impedance tomography using the extended kalman filter". *IEEE T Bio-Med Eng*, Vol. 51, pp. 72–81.

10. Responsibility notice

The author(s) is (are) the only responsible for the printed material included in this paper