Proceeding Paper

Comparing two Fitting Algorithms for Determining the Cole–Cole Parameters in Blood Glucose Problems †

Roberto Dima *, Giovanni Buonanno ⊗ and Raffaele Solimene ⊙

Department of Engineering, University of Campania, 81031 Aversa, Italy; giovanni.buonanno@unicampania.it (G.B.); raffaele.solimene@unicampania.it (R.S.)
* Correspondence: roberto.dima@unicampania.it
† Presented at the 2nd International Electronic Conference on Applied Sciences, 15–31 October 2021; Available online: https://asec2021.sciforum.net/.

Abstract: This paper addresses the nonlinear inverse problem of estimating the parameters of the Cole–Cole model used to describe the behavior of the complex permittivity of blood samples. Such a model provides an efficient and accurate representation of biological tissues in the entire frequency band considered and reduces the complexity of the experimental data to a few parameters. In this way, it is possible to extract a “synthetic view” of the dielectric properties of tissues in such a way that more information on the glucose concentration can be derived, in addition to the resonance peak or phase shift. In order to perform the fitting of the Cole–Cole model, two different algorithms were used and compared: the Levenberg–Marquardt and the variable projection algorithms. Synthetic data present in the literature were used to evaluate the performances obtainable with these methods. In particular, Monte Carlo analysis was used in order to evaluate the accuracy and the precision that these two methods provide in the process of estimating the parameters involved, with respect to the starting points of the parameters. The results obtained showed that the variable projection algorithm always outperformed the Levenberg–Marquardt one, although the former has a greater computational burden than the latter.

Keywords: glucose measurement; Cole–Cole model; Levenberg–Marquardt algorithm; variable projection algorithm; blood dielectric properties; nonlinear fitting problem

1. Introduction

Diabetes is a metabolic disorder that afflicts millions of people in the world [1]. It degrades the cell’s ability to absorb glucose from the bloodstream because of the improper regulation of the insulin hormone. For this reason, great efforts have been dedicated to the development of non-invasive glucose monitoring devices, which may considerably improve the quality of life for diabetics [2].

This work is particularly concerned with microwave sensor technology that relies on the change in the dielectric and conductivity properties of blood plasma as a function of the glucose concentration in order to track such a change.

In this framework, developing accurate and precise fitting methods for blood models, at different glucose concentrations, is essential for the development of robust electromagnetic (EM)-based techniques that could be employed for non-invasive, continuous glucose monitoring. Indeed, accurate electromagnetic tissue modeling is of paramount importance since it affects the simulation stage required for sensor design [3]. Moreover, extracting a “synthetic view” (in terms of a few parameters) of the sensor response data is essential for analyzing patterns and possibly extracting more information, besides the resonance peak or phase shift, about glucose concentration.

In this paper, the aim was just to address the fitting problem. More in detail, starting from the dielectric spectrum, which is assumed known over a certain number of frequencies, we aimed at estimating the parameters of a single-pole Cole–Cole model. As is well
known, this entails solving a nonlinear inverse problem, which here was addressed by
two different methods: the classical Levenberg–Marquardt method [4, 5] and the variable
projection algorithm [6]. We evaluated how sensitive the two methods are with respect to
the starting points of the parameters and with what accuracy and precision these param-
eters can be estimated. In order to check the two methods, we first generated synthetic
relative permittivities by employing a single-pole Cole–Cole model, using data from the
literature [7] as true values for its parameters, and then solved an inverse problem in order
to trace these values by resorting to the two aforesaid methods.

Although higher-order models could perform better, we considered a first-order model
to perform the comparison in the simplest possible case.

2. Methods

2.1. Cole–Cole Model

The Cole–Cole model [8] is widely used to describe the complex relative permittivity
of biological tissues, \( \varepsilon_r(\omega) = \varepsilon(\omega)/\varepsilon_0 \), and its equation is:

\[
\varepsilon_r(\omega) = \varepsilon_\infty + \sum_{n=1}^{N} \frac{\varepsilon_{s\infty} - \varepsilon_\infty}{1 + (j\omega\tau_n)^{1-\alpha_n}} - \frac{\sigma}{\omega\varepsilon_0}
\]  

(1)

in which \( N \) is the number of poles and thus the order of the model, \( \varepsilon_\infty = \lim_{\omega \to \infty} \varepsilon_r(\omega) \) is
the permittivity at high frequencies, \( \sigma \) is the static ionic conductivity and \( \varepsilon_{s\infty} = \lim_{\omega \to 0} \varepsilon_r(\omega) \),
and \( \tau_n \) and \( \alpha_n \) are the static permittivity, the relaxation time constant, and the so-called
distribution parameter of the \( n \)-th addend of the summation, respectively. Such a model
incorporates the Debye model [9]. Indeed, the main difference between the Debye and the
Cole–Cole models is that the latter includes the exponent \( 1 - \alpha \), with \( 0 \leq \alpha \leq 1 \). When
the exponent becomes smaller, the relaxation time distribution becomes broader, i.e.,
the transition between low- and high-frequency values becomes wider and the peak on the
imaginary part of the spectrum also becomes wider.

The complexity of both the structure and composition of biological material is such
that the dispersion region of each pole may be broadened by multiple contributions to it.
The broadening of the dispersion could be empirically accounted for by using the Cole–Cole
model [10], which is expected to give more accurate dielectric spectrum curve fitting.

2.2. Curve Fitting Algorithms

Let be \( x \) the vector of model parameters and \( P \) its length and \( M \) the number of
frequency points for which the measures are taken. We define the data vector ( \( \mathbf{T} \) stands
for transposition) \( y = [y(\omega_1) \ldots y(\omega_m) \ldots y(\omega_M)]^\top \) in which the \( m \)-th component of
the vector \( y \) is the observed value \( y(\omega_m) \). Let also \( \varepsilon_r = [\varepsilon_r(\omega_1; x) \ldots \varepsilon_r(\omega_m; x) \ldots \varepsilon_r(\omega_M; x)]^\top \)
be the model vector, here given by Equation (1), with \( \varepsilon_r(\omega_m; x) \) being the estimation at \( \omega_m \).

Solving the least-squares problem means finding \( \hat{x} \) such that:

\[
\hat{x} = \arg \min_{x \in \mathbb{R}^P} \left\{ \frac{1}{2} \| \varepsilon_r(x) - y \|_2^2 \right\}
\]

(2)

in which the function to minimize, \( \Psi = \frac{1}{2} \| \varepsilon_r(x) - y \|_2^2 \), is the \( \ell_2 \) quadratic norm of the
misfit \( r = \varepsilon_r(x) - y \), which is a nonlinear function such that \( r : \mathbb{R}^P \rightarrow \mathbb{R}^M \) with \( P \ll M \).

We addressed the nonlinear fitting problem with two methods: the Levenberg–
Marquardt algorithm (LMA) and the variable projection algorithm (VPA).

The Levenberg–Marquardt algorithm [4, 5] acts more as a gradient-descent method
when the parameters are far from their optimal value and acts more as the Gauss–Newton
method when the parameters are close to their optimal value [11]. The equation for the
step \( h \) at the \( k \)-th iteration is:

\[
\left( J(x_k)^\top J(x_k) + \lambda_k I \right) h = -J(x_k)^\top f(x_k)
\]

(3)
where $J$ is the Jacobian of $f$ and $\lambda_k$ is the damping parameter. It controls both the magnitude and direction of $h$, and it is chosen at each iteration. It can be shown [5] that, at each iteration, Equation (3) solves the minimization problem over a reduced set of admissible solutions, i.e., those that satisfy $\|h\| \leq R(\lambda)$, limiting the correction step to within a region near $x_k$. The radius of the trust region $R = R(\lambda)$ is a strictly decreasing function with $\lim_{\lambda \to 0} R(\lambda) = 0$. When $\lambda_k = 0$, the step $h$ is identical to that of the Gauss–Newton method, i.e., the same direction and maximum magnitude. As $\lambda \to \infty$, $h$ tends towards the steepest descent direction, with the magnitude tending towards 0.

Based on the above, we inferred the qualitative update rule for $\lambda_{k+1}$: If $\Psi(x_k + h) < \Psi(x_k)$, then the quadratic approximation works well, and we can extend the trust region, i.e., it will be $\lambda_{k+1} < \lambda_k$. Otherwise, the step is unsuccessful and we reduce the trust region, i.e., it will be $\lambda_{k+1} > \lambda_k$; in this way, the next step tends towards the negative gradient method, and a lower value of $\Psi$ is more likely to be found.

The MATLAB implementation was used, in particular the lsqcurvefit function with the Levenberg–Marquardt option.

The variable projection algorithm [6] is a method used to solve separable nonlinear least-squares problems. The least-squares problem is said to be separable when the model parameters can be separated into two sets of parameters, one that enters linearly into the model, $c = [c_1, \ldots, c_k]$, and another set of parameters that enter the model nonlinearly, $a = [a_1, \ldots, a_l]$, so that $x = [c, a]$. For each observation $y_m$ of a separable nonlinear least-squares problem, the model is a linear combination of nonlinear functions that depend on nonlinear parameters, and the model function can be written as $\varepsilon_r(\omega) = \sum_{j=1}^{k} c_j \phi_j(\omega; a)$.

The functional $\Psi$ is written in terms of residual vector $r$ as:

$$\Psi(a, c) = \frac{1}{2} \| y - \Phi(a)c \|^2$$

(4)

in which the columns of the matrix $\Phi$ are the nonlinear functions $\phi_j(\omega; a)$. The linear parameters $c$ could be obtained from the knowledge of $a$, by solving the linear least-squares problem:

$$c = \Phi(a)^{\dagger} y$$

(5)

which stands for the minimum-norm solution of the linear least-squares problem for fixed $a$, where $\Phi(a)^{\dagger}$ is the Moore–Penrose generalized inverse of $\Phi(a)$. By replacing this in Equation (4), we obtain the variable projection functional:

$$\Psi_{VP}(a) = \frac{1}{2} \| y - \Phi(a)^{\dagger} y \|^2$$

(6)

The variable projection algorithm consists of two steps: first minimizing Equation (6) with an iterative nonlinear method and then using the optimal value found for $a$ to solve for $c$ in Equation (5) [12]. The principal advantage is that the iterative nonlinear algorithm used to solve the first minimization problem works in a reduced space and less initial guesses are necessary. A robust implementation in MATLAB, called VARPRO [13], was adapted and used to deal with complex-valued problems, choosing the Levenberg–Marquardt option for the solution of Equation (6).

2.3. Numerical Simulations

The generation of the synthetic complex relative permittivity of blood plasma relied on quadratic fits to glucose-dependent Cole–Cole parameters reported in [7]; in particular, we considered two different concentrations, 100 mg/dL and 250 mg/dL, one normal and the other of severe diabetes, respectively, in accordance with the diagnostic criteria in [1]. The data vector consisted of $M = 1000$ points in the frequency range 500 MHz–20 GHz.

In gradient-like algorithms (such as those used in this paper), the choice of the initial point is a crucial factor for the convergence of the procedure. For the single-pole model case, it is fairly easy to exploit the physical meaning of the parameters to infer an initial...
estimate. However, since the noise can invalidate the initial estimate, we propose to study the robustness of the two algorithms with respect to the initial point. To this end, a Monte Carlo analysis was performed, iteratively evaluating the deterministic algorithms using a set of N = 1000 uniformly distributed random initial points arranged in a 5D hypercube of the parameter space in order to statistically characterize the results. Each side of the hypercube represents an interval containing the range of variation of each parameter for the glucose concentrations considered.

We ran simulations on a machine with Intel i9-10850K (10 physical cores), 32 GB RAM, and Ubuntu 21.04, and we took advantage of the Parallel Computing Toolbox, using the parfor loop for running the 1000 simulations.

The intervals for generating the random initial value for each parameter (of the Cole–Cole model) were chosen from the data tabulated in [7]. In particular, the widths of these intervals were the same for each glucose concentration and were: [1, 5] for $\varepsilon_{\text{true}}$, [1 × 10^{-14}, 1 × 10^{-11}] for $\varepsilon_{\text{true}}$, [0.1 - 1 × 10^{-9}, 0.1 + 1 × 10^{-9}] for $\tau$, [0.1, 5] for $\alpha$, and [0, 5] for $\sigma$. These intervals are relatively large compared to the values taken from [7] in order to test the two algorithms in sufficiently stressful situations. Only the range of variation of $\alpha$ is extremely small because the model used in [7] practically fixes it a priori to 0.1. Obviously, it must be taken into account that the VPA requires only generating the values for $\tau$ and $\alpha$.

For an initial blind qualitative assessment, we established evaluation intervals (the same for generations) for the estimated parameters so that we could assert that a reconstruction was “good” if it falls within these ranges, and “wrong” otherwise.

Now, let $\hat{x}^{(i)} = [\hat{\varepsilon}_{\text{true}}, \hat{\tau}, \hat{\alpha}, \hat{\sigma}]$ be the vector of the parameter estimates returned by the two algorithms at the i-th simulation, and let $\hat{\varepsilon}^{(i)}$ denote one of its five elements. Moreover, let $\langle x \rangle = (1/N_{\text{sim}}) \sum_{i=1}^{N_{\text{sim}}} \hat{x}_i$ and $\sigma_x = \sqrt{[1/(N_{\text{sim}} - 1)] \sum_{i=1}^{N_{\text{sim}}} [\hat{x}_i - \langle x \rangle]^2}$ be the sample mean and standard deviation, respectively, calculated for each parameter.

For a quantitative evaluation of the performance of the two algorithms, we then define multiple figures of merit for statistically characterizing the results of Monte Carlo analysis. For each parameter, Equations (7) and (8) define measures of accuracy and precision, respectively, defined over the entire set of reconstructions. However, such measures can be greatly affected by estimates that are very far from the true value, $x_{\text{true}}$, where the latter represents one of the five elements of the vector of reference values $x_{\text{true}} = [\hat{\varepsilon}_{\text{true}}, \hat{\tau}_{\text{true}}, \hat{\alpha}_{\text{true}}, \hat{\sigma}_{\text{true}}]$. For this reason, we also introduced $A_{\text{cut}}$ and $P_{\text{cut}}$ in Equations (9) and (10) in order to define the accuracy and precision measures, respectively, which instead dampen the effect of the above isolated events. The subscript cut means that they were calculated on a subset obtained by eliminating $\zeta\%$ of reconstructions with lower values and $\zeta\%$ of reconstructions with higher values, in which $0 < \zeta < 50$.

$$A = \left| \frac{x_{\text{true}} - \langle x \rangle}{x_{\text{true}}} \right| \times 100\%$$  
(7)

$$P = \frac{\sigma_x}{\langle x \rangle} \times 100\%$$  
(8)

$$A_{\text{cut}} = \left| \frac{x_{\text{true}} - \langle x \rangle_{\text{cut}}}{x_{\text{true}}} \right| \times 100\%$$  
(9)

$$P_{\text{cut}} = \frac{\sigma_{x_{\text{cut}}}}{\langle x \rangle_{\text{cut}}} \times 100\%$$  
(10)

3. Results

We conducted many numerical simulations by widening more and more the generation intervals. In this paper, we report the case where the generation intervals are very large except for the $\alpha$ interval, due to the above explanation. In all these experiments, following the qualitative criterion mentioned above, the VPA always provided good estimations, while the same was not true for the LMA. In particular, in the case considered, the LMA provided the wrong estimates in about 260 simulations out of 1000, for each of the two
glucose concentrations, while no wrong estimate was returned by the VPA. Here, by the “wrong” estimate, we mean that at least one component of the parameter vector $x$ had a value that was outside its generation range. Graphical representations of those qualitative results are provided in Figure 1.

**Figure 1.** Graphical representations of the convergence of the LMA and VPA for 1000 simulations and 2 different glucose concentrations: (a) the LMA for 100 mg/dL, (b) the LMA for 250 mg/dL, (c) the VPA for 100 mg/dL, and (d) the VPA for 250 mg/dL.

Consistent with the qualitative results, considering the whole set of 1000 estimates, the VPA exhibited excellent accuracy and precision, while this was not the case for LMA, as can be observed in Tables 1 and 2. For this reason, for the VPA only, the results calculated by means of Equations (7) and (8) are reported, whilst for the LMA, the results derived from Equations (9) and (10), obtained by cutting 25% of the lowest values and 25% of the highest values, were also considered.

Regarding execution times, the LMA took 7 s to finish 1000 simulations, while the VPA took around 170 s.

**Table 1.** Figures of merit (in %) of the two algorithm in which the glucose concentration is 100 mg/dL.

| $x_{\text{true}}$ | $\varepsilon_\infty$ | $\varepsilon_s$ | $\tau$ (s) | $\alpha$ | $\sigma$ (S/m) |
|-----------------|------------------|----------------|----------|--------|---------------|
| **LMA**         |                  |                |          |        |               |
| $A, A_{\text{cut,25%}}$ | $2.043 \times 10^5$ | $5.34 \times 10^4$ | $1.383 \times 10^5$ | $4.161 \times 10^5$ | $1.921$ |
| $\beta, \beta_{\text{cut,25%}}$ | $2.257 \times 10^{-7}$ | $1.799 \times 10^{-10}$ | $1.041 \times 10^{-9}$ | $3.237 \times 10^{-9}$ | $6.762 \times 10^{-10}$ |
| **VPA**         |                  |                |          |        |               |
| $A, A_{\text{cut,25%}}$ | $1.305 \times 10^{-7}$ | $4.024 \times 10^{-10}$ | $6.201 \times 10^{-9}$ | $2.129 \times 10^{-8}$ | $4.876 \times 10^{-10}$ |
| $\beta, \beta_{\text{cut,25%}}$ | $3.569 \times 10^{-7}$ | $1.13 \times 10^{-9}$ | $1.694 \times 10^{-8}$ | $5.877 \times 10^{-8}$ | $1.357 \times 10^{-9}$ |

**Table 2.** Figures of merit (in %) of the two algorithm in which the glucose concentration is 250 mg/dL.

| $x_{\text{true}}$ | $\varepsilon_\infty$ | $\varepsilon_s$ | $\tau$ (s) | $\alpha$ | $\sigma$ (S/m) |
|-----------------|------------------|----------------|----------|--------|---------------|
| **LMA**         |                  |                |          |        |               |
| $A, A_{\text{cut,25%}}$ | $1.344 \times 10^5$ | $3.58 \times 10^6$ | $2.692 \times 10^5$ | $2.468 \times 10^7$ | $1.995$ |
| $\beta, \beta_{\text{cut,25%}}$ | $2.248 \times 10^{-7}$ | $1.699 \times 10^{-10}$ | $1.101 \times 10^{-9}$ | $3.316 \times 10^{-9}$ | $7.198 \times 10^{-10}$ |
| **VPA**         |                  |                |          |        |               |
| $A, A_{\text{cut,25%}}$ | $1.322 \times 10^{-7}$ | $4.107 \times 10^{-10}$ | $6.306 \times 10^{-9}$ | $2.17 \times 10^{-8}$ | $5.029 \times 10^{-10}$ |
| $\beta, \beta_{\text{cut,25%}}$ | $3.595 \times 10^{-7}$ | $1.153 \times 10^{-9}$ | $1.712 \times 10^{-8}$ | $5.965 \times 10^{-8}$ | $1.396 \times 10^{-9}$ |

4. Discussion

In this paper, we faced the problem of fitting the dielectric spectrum of a blood sample in order to estimate the parameters of the single-pole Cole–Cole model. In particular, we
compared the performance of two different algorithms, the LMA and VPA, in terms of the accuracy and precision with respect to the starting points of the parameters.

For the parameter range considered, the VPA outperformed the LMA in robustness with respect to the initial point of the algorithm. However, analyzing the figures of merit related to the LMA, it became clear that there were erroneous reconstructions so far from the true value such that they heavily deteriorated the (standard) accuracy and precision, while the cut versions did not suffer from this problem. In fact, once the erroneous ones were removed, in all other simulations, the algorithm converged to the true values.

On the other hand, the VPA gave these good results because less initial guesses were necessary and because the iterative nonlinear algorithm used to solve the first minimization problem works in a reduced space. The big disadvantage of the VPA, or at least of the implementation used in this paper, was the execution time. This was certainly due to the numerous singular-value decompositions (SVDs) that the algorithm calculates in its runtime.

The results are promising, and the research will continue by evaluating the algorithms in increasingly realistic scenarios, including adding noise to the synthetic data.

Author Contributions: Conceptualization, R.S.; methodology, R.D. and G.B.; software, R.D.; validation, R.D. and G.B.; formal analysis, R.S., R.D. and G.B.; investigation, R.D. and G.B.; resources, R.S.; data curation, R.D.; writing—original draft preparation, R.D.; writing—review and editing, G.B. and R.D.; visualization, R.D. and G.B.; supervision, R.S. and G.B.; project administration, R.S.; funding acquisition, R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by MIUR, Project Grant PRIN 2017 “Microwave Biosensors: Enhanced Non-Invasive Methodology for Blood Glucose Monitoring”. The APC was funded by the above Project Grant PRIN 2017.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. IDF Diabetes Atlas, 9th ed.; Technical Report; International Diabetes Federation: Brussels, Belgium, 2019.
2. Lin, T. Non-Invasive Glucose Monitoring: A Review of Challenges and Recent Advances. *Curr. Trends Biomed. Eng. Biosci.* 2017, 6, 1–8. [CrossRef]
3. Costanzo, S.; Cioffi, V. Dielectric Models for the Accurate Design of Wearable Diabetes Sensors. In Proceedings of the 2019 23rd International Conference on Applied Electromagnetics and Communications (ICECOM), Dubrovnik, Croatia, 30 September–2 October 2019; pp. 1–3. [CrossRef]
4. Levenberg, K. A Method for the Solution of Certain Non-Linear Problems in Least Squares. *Q. Appl. Math.* 1944, 2, 164–168. [CrossRef]
5. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Ind. Appl. Math.* 1963, 11, 431–441. [CrossRef]
6. Golub, G.H.; Pereyra, V. The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate. *SIAM J. Numer. Anal.* 1973, 10, 413–432. [CrossRef]
7. Karacolak, T.; Moreland, E.C.; Topsalak, E. Cole–Cole Model for Glucose-Dependent Dielectric Properties of Blood Plasma for Continuous Glucose Monitoring. *Microw. Opt. Technol. Lett.* 2013, 55, 1160–1164. [CrossRef]
8. Cole, K.S.; Cole, R.H. Dispersion and Absorption in Dielectrics I. Alternating Current Characteristics. *J. Chem. Phys.* 1941, 9, 341–351. [CrossRef]
9. Debye, P. *Polar Molecules*; The Chemical Catalog Company, Inc.: New York, NY, USA, 1929.
10. Gabriel, S.; Lau, R.W.; Gabriel, C. The Dielectric Properties of Biological Tissues: III. Parametric Models for the Dielectric Spectrum of Tissues. *Phys. Med. Biol.* 1996, 41, 2271–2293. [CrossRef] [PubMed]
11. Gavin, H.P. The Levenberg–Marquardt Algorithm for Nonlinear Least Squares Curve-Fitting Problems; Duke University: Durham, NC, USA, 2019; p. 19. [CrossRef]
12. Golub, G.; Pereyra, V. Separable Nonlinear Least Squares: The Variable Projection Method and Its Applications. *Inverse Probl.* 2003, 19, R1–R26. [CrossRef]
13. O’Leary, D.P.; Rust, B.W. Variable Projection for Nonlinear Least Squares Problems. *Comput. Optim. Appl.* 2013, 54, 579–593. [CrossRef]