19th Symposium IMEKO TC 4 Symposium and 17th IWADC Workshop
Advances in Instrumentation and Sensors Interoperability
July 18-19, 2013, Barcelona, Spain

Improving the convergence of gene expression programming in impedance spectroscopy

Pedro M. Ramos[1], Fernando M. Janeiro[2]

[1] Instituto de Telecomunicações, Instituto Superior Técnico, DEEC, UTL, Av. Rovisco Pais 1, Lisboa, Portugal, pedro.m.ramos@ist.utl.pt
[2] Instituto de Telecomunicações, Universidade de Évora, Rua Romão Ramalho 59, Évora, Portugal, fmtj@uevora.pt

Abstract- This paper describes the latest major improvement made to gene expression programming (GEP) for use in impedance spectroscopy. This change consists on systematically analysing the fittest element of each population of the GEP to identify circuit components that are useless in the sense that they do not significantly contribute to the impedance response at the analysed frequency range. These components are then removed from the circuit making it less complex. The performance of the proposed improvements is analyzed on both simulated and measured impedance frequency responses.

I. Introduction

The broad range of applications of impedance spectroscopy [1] has led to the continuous development of its methods and techniques. In the soft computing area, the developments have focused on automatically finding suitable equivalent circuits that correctly model the measured impedance responses. This has been achieved for example with Gene Expression Programming (GEP) [2] that was originally developed to overcome the shortcomings of Genetic Programming. A tree structure was proposed in [3] that allowed the use of GEP on electric circuits. In GEP, a set of randomly generated circuit topologies is checked against the impedance frequency response to be modelled, and then evolved according to rules that have origin in genetic theory. GEP has been coupled to Cultural Algorithms [4] in an attempt to automatically model electrochemical phenomena.
Another approach has consisted on using GEP together with Genetic Algorithms (GA) which has proven to be a useful and efficient tool in impedance spectroscopy [5] and sensor modelling [6, 7] when the circuit model is unknown. Nevertheless, the complexity of the problem has limited the performance of the GEP+GA solution and its applicability. Thus, attempts have been made at improving the GEP+GA algorithm performance. The use of compound components has led to smaller genes and higher circuit diversity in the GEP search procedure [7]. Simultaneously, an automatic circuit simplification routine has been developed that has introduced improvements in the algorithms total execution time and on the amount of components needed to correctly model the measured impedance. However, the long term objective has been to improve the GEP+GA convergence to the simplest circuit that correctly models the measurements. To this end, a new approach that analyses the fittest element of each GEP population and attempts to remove circuit components that do not contribute to the impedance response on the frequency range under study, has been developed and deployed.
In this paper, these new improvements to the GEP implementation for impedance spectroscopy are described. Its performance is analysed on simulated impedance frequency responses. Finally, to evaluate its suitability to real world conditions, its performance has been analysed on measurement data.
Section II details the improvements proposed in this paper. Section III includes a performance analysis of the new algorithm when compared with previous versions from [6]. Section IV includes measurement results that demonstrate the usefulness of this method. In section V, the main conclusions of this work are listed.

II. Improvements to gene expression programming

In gene expression programming, there is a fixed size population of circuits which are evaluated using GA to estimate the fitness value of each circuit (this fitness value is the absolute relative difference between the impedance response of the circuit under evaluation and the measured impedance response [6]). In previous versions, the use of compound components and the inclusion of a circuit simplification routine based on the analysis of the circuit topology (for example, joining identical components in series or in parallel into a single component) has proven to be very effective in the reduction of the gene size and decrease the relative execution time of the full GEP+GA routine.
The improvement proposed in this paper is that, in addition to the previous improvements, at the end of each GEP generation, the best circuit is further analysed by selectively changing (or eliminating) the gene elements to see if they actually change the circuit impedance response within the measured frequency range.

In Figure 1, a starting circuit example is shown to illustrate this circuit optimization process. For this particular circuit, three tree elements are analysed. Element #1 is a RL series circuit. To analyse the influence of these two components on the circuit impedance (within the measured frequency impedance range), three tests are conducted. In the first test, the inductance is removed and only the resistance remains (Figure 2a). In the second test it is the resistance that is removed and only the inductance remains (Figure 2b). The third and final test eliminates completely the RL series circuit and what remains is the left-hand side of the original tree (Figure 2c). These three new circuits are analysed to determine if the circuit fitness is at least as good as the cost function value of the original circuit (in fact, to enable circuit simplification, an increase of 5% is accepted to reduce any over fitting that might occur). If any of these three circuits complies with these conditions, the original circuit is eliminated and the new best circuit takes its place. At this point, instead of resuming the original GEP+GA algorithm, the circuit optimization routine restarts now with the simplified circuit. This is done to attempt to include further optimizations without repeating an extra, eventually unnecessary, GEP+GA iteration. Notice that, whether the GEP population is 20 or 50 (typical values), this step involves only the assessment of a few number of new circuits (basic circuit simplification is executed and then GA is used to estimate the component values and obtain the cost function value). Thus the overhead is not necessarily the one needed for a full GEP+GA iteration and is instead focused on the improvement of the best GEP+GA circuit of the previous iteration.

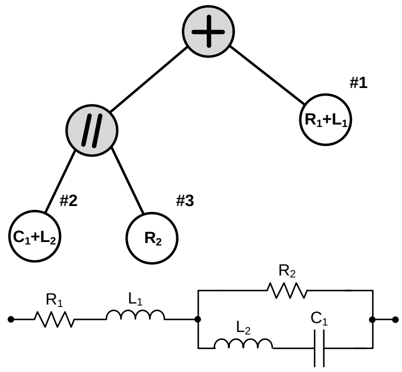

Figure 1. Example of tree structure and corresponding impedance circuit.

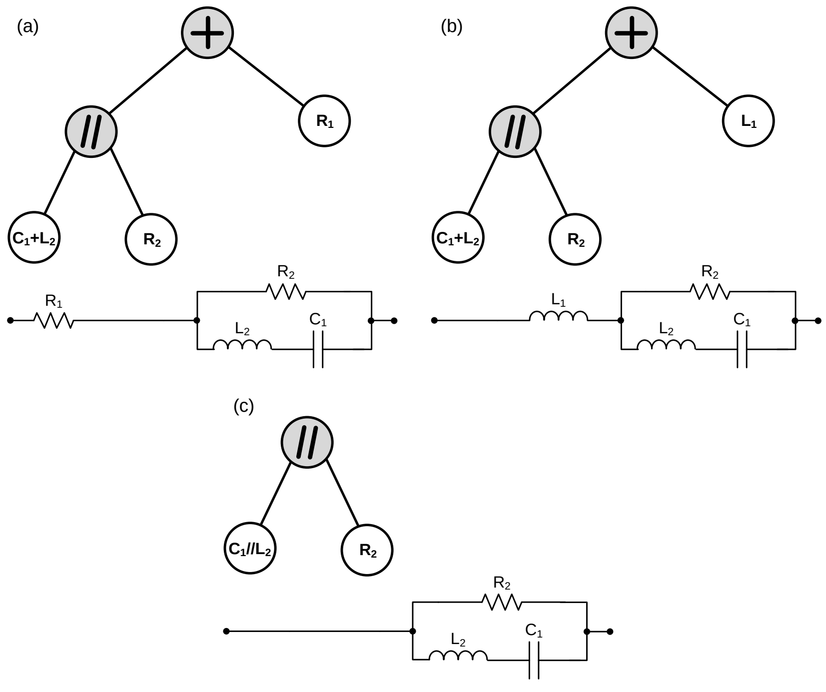

Figure 2. Three tested simplified circuits originated from the circuit of Figure 1 when element #1 is changed. (a) corresponds to eliminating L_1, (b) to eliminating R_1 and (c) to completely eliminate element #1.

ISBN-10: 84-616-5438-2 | ISBN-13: 978-84-616-5438-3

If none of the circuits from Figure 2, proves to be a better fit to the measured impedance values than the original circuit of Figure 1, element #2 is now considered as an option for optimization. Since it is a LC series, three options are also analysed. When element #2 is replaced with a capacitance, it combines with R from element #3 and forms a RC parallel circuit (Figure 3a). On the other hand, when element #2 is replaced by a single inductance, it combines with element #3 for from a RL parallel element (Figure 3b). The complete elimination of element #2 results in a simple RL series circuit since the R in element #3 is in series with the RL series combination of element #1 (Figure 3c). These three circuits are analysed as in the previous circuit simplification attempt.

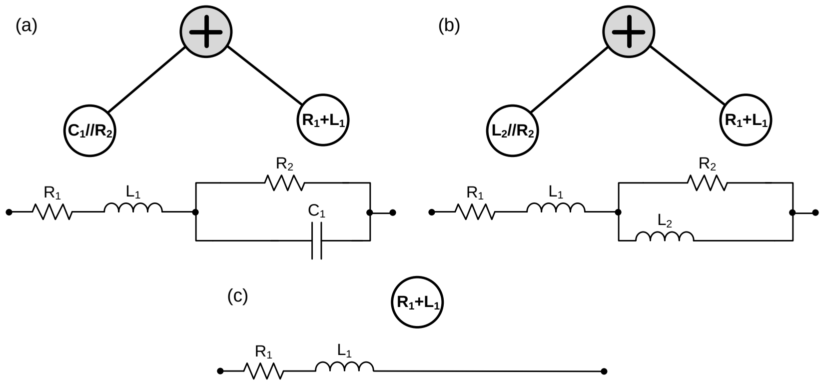

Figure 3. Three tested simplified circuits originated from the circuit of Figure 1 when element #2 is changed. (a) corresponds to eliminating L_2, (b) to eliminating C_1 and (c) to completely eliminate element #2 which simplifies to an RL series circuit.

If also, none of these three new circuits is able to improve the cost function value (thus indicating that they would be a better fit than the circuit of Figure 1), the last element (#3) is now considered. Since this element is a simple component, the only tested circuit corresponds to its elimination. The resulting circuit is a LC series (element #1) in series with a RL series (element #2). This circuit is simplified to an RLC series since the inductances can be replaced by a single one. The resulting circuit is shown in Figure 4.

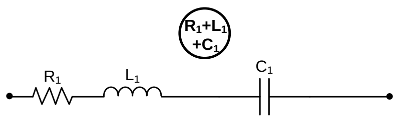

Figure 4. Unique tested simplified circuit originated from the circuit of Figure 1 when element #3 is eliminated which simply corresponds to an RLC series circuit.

In the end of the complete optimization routine, the cost function value of the best circuit is compared with the initially defined threshold to detect convergence of the GEP+GA algorithm. If the new cost function value is below that threshold, the complete GEP+GA (with the circuit optimization) iteration has converged and the circuit is found. If the cost function value is above the threshold, a new GEP iteration is started.

III. Improved results

To compare the performance of the improved implementation, its results were compared with the three previous versions: Version 1 has no circuit simplification and no combined circuit elements. Version 2 has circuit simplification while Version 3 has circuit simplification and combined circuit elements. The test circuit is a series circuit between a series RL and a parallel RLC as used in [6] for the three previous versions. This impedance is simulated with 1000 impedance frequency values in the 100 Hz to 10 kHz with 100 Hz steps. The threshold value to detect GEP+GA convergence is set to 2×10^{-6}. To ensure identical conditions to the measurement situation, measurement uncertainties were included where random values are added to the impedance magnitude and phase according to Gaussian distributions with standard deviation of 0.08% and 0.05° respectively. These values correspond to the uncertainty values of a commercial impedance measurement device from HIOKI [8].

ISBN-10: 84-616-5438-2 | ISBN-13: 978-84-616-5438-3

In Figure 5 the convergence percentages are shown for the four versions. It can be seen that the new version registers full convergence but most significantly of all, the convergence to the correct circuit has increased from around 15% to around 57%. Additionally, when compared with Version 3 (the previous best), the average number of iterations was reduced from 6.9 to 3.8 and the average gene size was also reduced from 5.7 to 4.3. Version 4 is slower (average 123.0 s/run) than Version 3 (average 85.8 s/run).

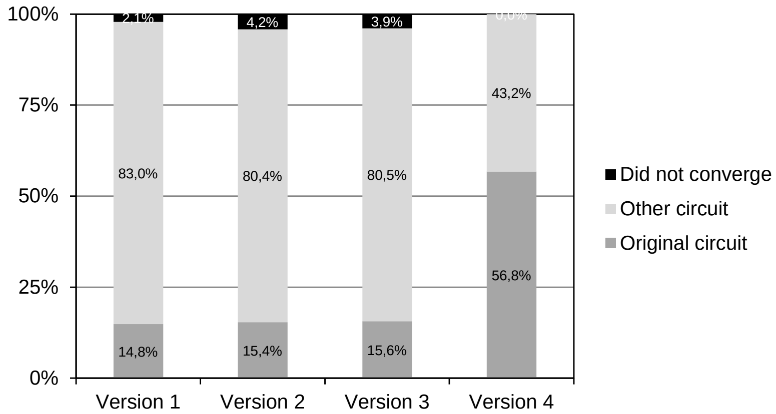

Figure 5. Comparison of the convergence percentage between the previous three versions and the presented improved version with circuit optimization (Version 4).

Figure 6 depicts the gene size obtained for the new Version 4 and also for Version 3 (the previous version without the circuit optimization stage proposed in this paper). The advantage of Version 4 is clearly seen in the percentage of occurrence of the correct gene size which is 3. The inclusion of the circuit optimization step enabled the elimination of the additional components introduced by GEP along the multiple iterations that, in the end, do not affect the impedance values within the measured frequency range. After these components are removed, the normal circuit simplification can, in many more situations, converge to the correct circuit.

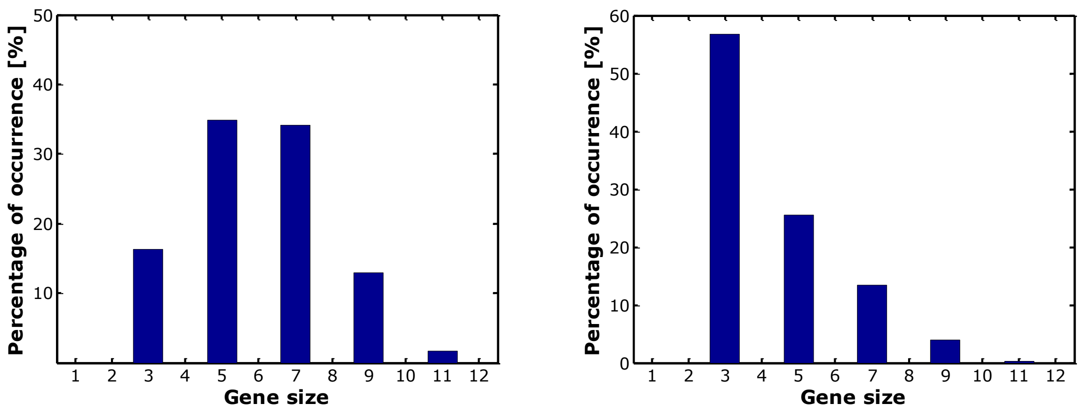

Figure 6. Percentage of occurrence of the gene size for 5000 runs of the GEP+GA algorithm for Version 3 (left) and Version 4 (right).

Although the improvements described so far, clearly demonstrate the evolution of the algorithm and that, therefore, its usefulness is much broader, there are still some aspects that can be improved. One of the most important aspects is related to the eventual reduction of over-fitting situations (where the algorithm can add additional components just to reduce the fitting error of one or two frequency impedance measurements). This issue is more relevant when there are few impedance measurements and quite spaced. So, in order to mitigate this effect, the circuit optimization routine was further improved by ensuring that the selection of the optimized circuit not only depends on the circuit with the lowest fitting error but is also based on the circuit the minimum amount of components that still complies with the maximum fitting error initially defined as the stop criteria to detect convergence of the GEP+GA algorithm.

ISBN-10: 84-616-5438-2 | ISBN-13: 978-84-616-5438-3

IV. Measurement Results

In this section, measurement results to demonstrate the applicability and effectiveness of the proposed algorithm are presented. The circuit represented in Figure 7 was measured using a DSP based impedance measurement instrument which is a development of the system presented in [9]. The system includes two analog to digital converters that measure the voltage across the unknown impedance and across a reference impedance. The current in both impedances is the same, and therefore with the use of sine-fitting algorithms, in particular two-channel tailored systems as the one presented in [10], it is possible to estimate the parameters that describe the analytical equations of both sine signals and from them to estimate the impedance magnitude and phase. The impedance represented in Figure 7 is measured at 50 different frequencies from 100 Hz up to 5 kHz with 100 Hz spacing. This particular impedance is characterized by a resonance near 1 kHz (caused by the RLC parallel) and a high-value impedance at lower frequencies due to the capacitor in series. In order to fully evaluate the performance of the proposed algorithm, 1000 sweeps were measured.

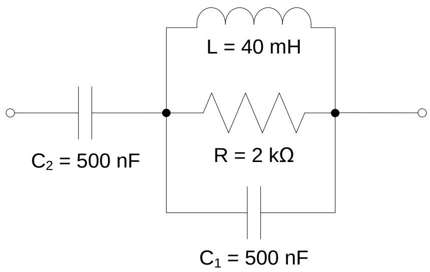

Figure 7. Measured impedance from 100 Hz up to 5 kHz with 100 Hz spacing.

In Figure 8, the results obtained with the fully improved algorithm are shown. The percentage of occurrence of estimation of the correct circuit has improved to 84.7% while the number of iterations is almost always 1. The experimental relative standard deviation for R is 0.11%, for C1 is 0.06%, for L is 0.06% and for C2 is 0.05%.

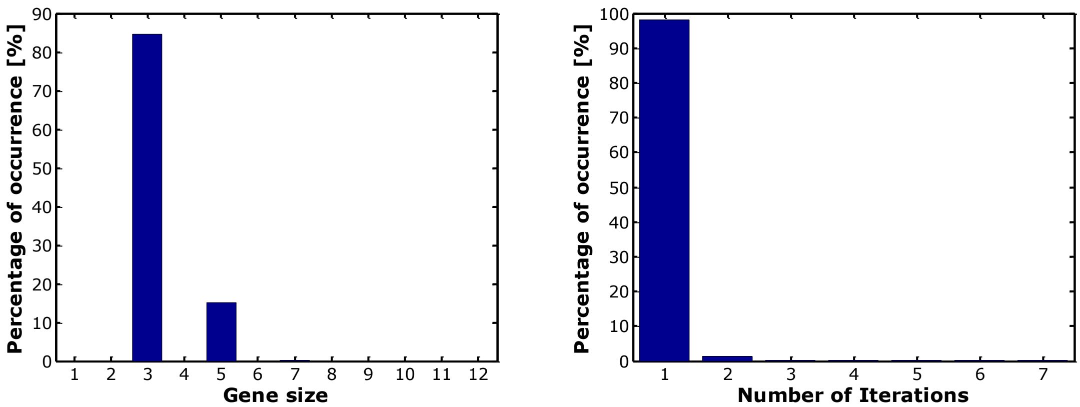

Figure 8. Percentage of occurance of the gene size and number of iterations. These results were obtained from the 1000 repetitions.

V. Conclusions

The latest improvements to the GEP+GA algorithm for impedance spectroscopy have been presented. The main development consists on the use of a circuit optimization routine that analyses the best circuit of each GEP+GA iteration and determines the usefulness of every component in the circuit. The motivation for this additional step is that, from the analysis of previous versions of the GEP+GA algorithm, most of the times, when it fails to converge to the correct circuit it is because additional components are added by the algorithm. Further analysis revealed that these components do not significantly change the circuit impedance within the frequency range at which the impedance was measured. Since these components cannot be removed by the circuit simplification step which is strictly based on absolute circuit equivalent rules, the circuit optimization routine was developed.

ISBN-10: 84-616-5438-2 | ISBN-13: 978-84-616-5438-3

The presented results clearly demonstrate that the proposed change has substantially increased the overall performance of the algorithm when compared with previous versions. Tests have shown that in both simulated and measured impedance frequency responses, the convergence of the GEP+GA algorithm, to the simplest correct circuit has dramatically increased when compared to the previous state of the art.

References

[1] E. Barsoukov, J. Macdonald, Impedance Spectroscopy Theory, Experiment, and Applications, Wiley Interscience, 2005.

[2] C. Ferreira. Gene Expression programming: mathematical modeling by an artificial intelligence: Springer-Verlag, 2006.

[3] J. Yu, H. Cao, and Y. He, “A new tree structure code for equivalent circuit and evolutionary estimation of parameters”, Chemometrics and Intelligent Laboratory Systems, vol. 85, pp. 27-39, 2007.

[4] P. Arpaia, “A cultural evolutionary programming approach to automatic analytical modeling of electrochemical phenomena through impedance spectroscopy”, Meas. Sci. Technol., vol. 20, pp. 1-10, 2009.

[5] Fernando M. Janeiro, Pedro M. Ramos, “Gene expression programming and genetic algorithms in impedance circuit identification”, ACTA IMEKO, vol. 1, n.º 1, pp. 19-25, 2012.

[6] P. M. Ramos, F. M. Janeiro, “Gene expression programming for automatic circuit model identification in impedance spectroscopy: Performance evaluation”, Measurement, accept for publication in 2013.

[7] Fernando M. Janeiro, José Santos, Pedro M. Ramos, “Gene expression programming in sensor characterization: Numerical results and experimental validation”, IEEE Transactions on Instrumentation and Measurement, vol. 62, n.º 5, pp. 1373-1381, May 2013.

[8] 3522-50/3532-50 LCR HiTester, Hioki E. E. Corporation, 2001.

[9] José Santos, Pedro M. Ramos, “DSPIC-based impedance measuring instrument”, Metrology and Measurement Systems, vol. XVIII, n.º 2, pp. 185-198, 2011

[10] Pedro M. Ramos, A. Cruz Serra, “A new sine-fitting algorithm for accurate amplitude and phase measurements in two channel acquisition systems”, Measurement, Elsevier, vol. 41, n.º 2, pp. 135-143, Feb. 2008.

ISBN-10: 84-616-5438-2 | ISBN-13: 978-84-616-5438-3