Advanced parameter optimization in EPR spectroscopy

B. Filipič[1] & J. Štrancar[2]
[1]*Dept. of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia*
[2]*Dept. of Solid State Physics, Jožef Stefan Institute, Ljubljana, Slovenia*

Abstract

We present an automated optimization procedure to enhance the characterization of biological systems with electron paramagnetic resonance (EPR). EPR is a physical phenomenon exploited in EPR spectroscopy to detect changes in organisms caused by biologically active substances or resulting from pathological conditions. In the past, EPR spectral characteristics were interpreted manually, while nowadays numerical spectrum simulation makes it possible to obtain more reliable and biologically meaningful information about the inspected system. For this purpose, the parameters of the spectrum simulation model need to be tuned so that the resulting simulated spectrum matches with the experimentally obtained EPR spectrum. To search for appropriate parameter values, we have integrated the simulation model with an evolutionary optimization algorithm. Our initial results indicated that this approach alleviates the weaknesses of previously used human-navigated optimization techniques and saves much of the spectroscopist's time. In this paper we provide further empirical results on the robustness of this approach and its application in cell membrane characterization.

1 Introduction

Electron paramagnetic resonance (EPR), also known as electron magnetic resonance (EMR) or electron spin resonance (ESR), denotes the physical phenomenon of absorption of microwave radiation by paramagnetic molecules or ions exposed to an external magnetic field. This phenomenon is exploited in EPR spectroscopy, which is a nondestructive method suitable for inspecting biological systems in their complex environments. The only alteration required for the native system is the insertion of spin probes or some other stable radicals into the system. An important

property of EPR spectroscopy is its ability of detecting heterogeneity in inspected systems. Heterogeneity reflects in superimposed EPR spectra consisting of several spectral components, also called domains. They arise from various compartments of the inspected system that exhibit different physical characteristics. Heterogeneity comes into play in the final stage of EPR inspection where the EPR spectrum needs to be interpreted. In the past, spectrum interpretation was performed manually by measuring spectral peak characteristics and analyzing their relationships. However, experimental EPR spectra provide much more reliable and biologically meaningful information when characterized by computer-aided spectrum simulation. Using an appropriate biophysical simulation model, it is possible to resolve different spectral components indicating different anisotropy and/or dynamics of the spin probes in various compartments of the inspected system [1, 2].

Model-based EPR spectrum interpretation requires the parameters of the biophysical model to be tuned so that the simulated spectrum matches with the recorded spectrum. When solving this problem with traditional single-point optimization techniques, a spectroscopist needs to provide good starting parameter values and perform a sequence of algorithm runs. As this approach is time-consuming and requires active user participation, an automated optimization procedure would allow the spectroscopist to focus on experiments and obtain better insight into the inspected biological system.

Potential candidates for automating parameter optimization in EPR spectroscopy are evolutionary computation techniques [3] that are successful in finding near-optimal solutions in complex search spaces. We tested this approach by integrating a spectrum simulation model with an evolutionary algorithm. The integrated system was experimentally evaluated on synthetic EPR spectra [4], and the results were promising both with regards to the accuracy of the resulting spectral parameters and the time spent by a spectroscopist using this method. The optimization procedure was upgraded into a hybrid version by incorporating the downhill simplex algorithm. This improvement resulted in higher accuracy and faster convergence of the method [5]. This paper provides new results on the robustness of the method and its results in cell membrane characterization. Unlike in previous experiments, where synthetic spectra were used, this study is based on real data obtained in spectroscopic inspection of liposomes.

The paper presents EPR spectroscopy as a method for characterization of biological systems, describes the parameter optimization procedure and recent experimental results in cell membrane characterization, and concludes with a summary of the work done and plans for the future work.

2 EPR spectroscopy in characterization of biological systems

2.1 Background

EPR spectroscopy is a powerful technique for characterization of biological systems, particularly when combined with physiological experiments. Combining the two approaches, we can measure various physiological quantities, such as, for

example, biochemical response of a tissue, muscle contraction, cell culture survival, metabolic and antioxidant activity, etc. Additionally, we can determine EPR-based parameters including domain weights, membrane fluidity, polarity, concentrations and others. By correlating measured physiological and EPR quantities, both caused by the same external factor, we can connect macroscopic response with microscopic changes in the structure and dynamics of the inspected system. Such combined experiments were, for example, performed to correlate the membrane characterization of plasma membranes of erythrocytes, activities of various enzymes and receptors, and symptoms of various diseases, such as acute phase and cancer.

2.2 Spectrum simulation and spectral parameters

In this study we apply EPR spectroscopy in cell membrane characterization. The biophysical model used to numerically simulate the EPR spectra of spin labeled membranes relies on the so called motional-restricted fast-motion approximation [6]. The model presumes multi-domain structure of a membrane and takes into account fast and anisotropic rotational motion of molecules. These assumptions are based on proven lateral heterogeneity of cellular membranes, as well as on experimental setup which includes small spin-labeled molecules and physiological temperatures. Parameters of the model provide information about ordering, dynamics and the polarity at various locations in different membrane domains.

EPR spectrum simulation with motional-restricted fast-motion approximation model requires the following parameters to be set for each spectral domain:

- order parameter S,
- rotation correlation time τ_c,
- broadening constant W,
- polarity correction factors p_A and p_g, and
- weighting factor d.

The order parameter S is the most important spectral parameter which in a way characterizes the type of a domain. It indicates the anisotropy of the actual orientation distribution relative to the membrane normal vector. Higher order parameter value means more compact and ordered domain in the sense of the directions of the molecular axes.

The rotational correlation time τ_c characterizes rotational motions of the molecular conformational changes. Larger values of rotational correlation times indicate rotational motions with less collisions and changes in direction and speed of motion.

The additional broadening constant W includes the effect of the spin-spin interaction in non-ideal diluted samples and contains information about the diffusion constants and local concentrations of the spin probes in particular compartment of the inspected system. Larger value of W indicates the increase of the product of the diffusion constant and concentration.

On the other hand, both polarity correction factors p_A and p_g reflect changes in the electron density distributions. They are influenced by the polarity of the local environment of the spin probes or by any other source of the electric fields (for example a group of polar molecules, electric charge double layer, etc.).

Finally, the weighting factor d denotes the proportion of a domain in the resulting simulated spectrum.

In general, for an EPR spectrum with k domains, values of $6k - 1$ spectral parameters need to be determined. Since $\sum_{i=1}^{k} d_i = 1$, the weighting factor for k-th domain, d_k, is equal to $1 - \sum_{i=1}^{k-1} d_i$.

Given the values of spectral parameters, the calculation procedure produces a simulated EPR spectrum. Figure 1 shows an example of the experimental spectrum and Figure 2 the related three-domain simulated spectrum.
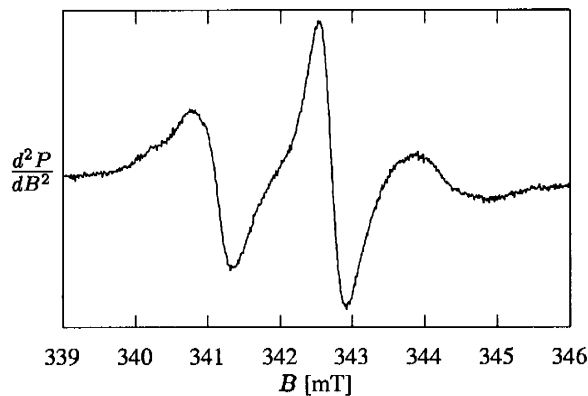

Figure 1: An experimental EPR spectrum: signal intensity (in arbitrary units) versus magnetic field density B.

The optimization task is to find parameter values such that the simulated spectrum fits with the experimental one as closely as possible. This task requires a powerful optimization tool since the number of the involved parameters is high and they are also partially correlated. The numerical spectrum simulation model should therefore be coupled with an optimization method that can provide the values of model parameters both with required accuracy and in acceptable period of time.

3 Optimization of spectral parameters

To automate the search for appropriate parameter values in EPR spectrum simulation, an iterative numerical optimization procedure was designed. It consists of an evolutionary algorithm (EA) which is hybridized by the downhill simplex method.
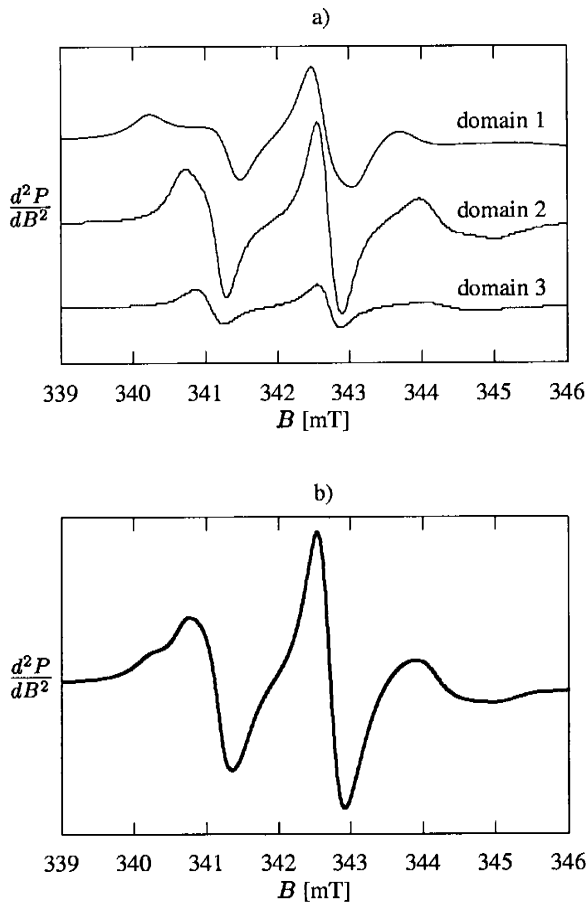
Figure 2: Numerical simulation of the experimental spectrum from Figure 1: a) spectral domains, b) the resulting spectrum.

Evolutionary algorithms [3] utilize the principles of natural evolution in computer problem solving. The underlying idea is to search for a good solution to a given problem through computer simulated evolution of candidate solutions. This procedure starts with a population of randomly created solutions and iteratively improves them by employing stochastic evolutionary mechanisms, such as survival of the fittest and exchange of genetic information. Evolutionary algorithms exhibit a number of advantages over traditional specialized methods and other stochastic algorithms. Besides the evaluation of candidate solutions, they require no additional information about the properties the search space. By processing populations of candidates, they are capable of providing alternative solutions to complex problems. They require low development costs and can be easily extended by incorporating elements of other search algorithms.

The EA for EPR spectral parameter optimization employs the real-valued vector representation of candidate solutions. The number of vector components depends on the number of the encountered spectral domains. For example, for a typical real problem with three spectral domains ($k = 3$) there are $6k - 1 = 17$ parameters to be optimized. The number of domains itself is not subject to optimization and needs to be provided by the user.

To evaluate a vector of parameter values, the algorithm activates the simulation model that generates a spectrum, and then assesses the goodness of fit of the simulated spectrum with the one recorded in the EPR experiment. The quality measure is the reduced χ^2, which is the sum of the squared residuals between the experimental and simulated spectra normalized by the squared standard deviation of the experimental points, σ, and by the number of points in the experimental spectrum, N:

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(y_i^{\text{exp}} - y_i^{\text{sim}})^2}{\sigma^2} . \tag{1}$$

The standard deviation σ is assessed numerically from the experimental points of those parts where the simulated spectrum derivatives are close to zero. This is usually at both ends of the spectrum. The smaller the χ^2 value, the better the setting of the spectral parameters. To be used with the evolutionary algorithm, this measure is transformed into an increasing fitness function by subtracting from a large positive constant.

The population of candidate parameter settings is iteratively improved by applying operators typical of evolutionary algorithms: selection, which probabilistically selects the best solutions from current population to participate in the next generation, crossover, which exchanges components between randomly selected pairs of parameter vectors, and mutation, which randomly alters vector components. For each spectral parameter the interval of possible values and step size are defined in advance according to physical limitations and user preferences.

As preliminary numerical experiments reported in [4] indicate that the evolutionary algorithm combined with other optimization techniques outperforms the poor variant of the algorithm, we incorporated the downhill simplex method of Nelder and Mead [7] into our algorithm. Downhill simplex is a deterministic multidimensional optimization method which, like evolutionary algorithms, requires only solution evaluations, not derivatives, and iteratively improves the solutions. Hybridization of the evolutionary algorithm was achieved by executing downhill simplex on probabilistically selected solutions during the evolutionary algorithm run, and by applying it to the to members of the final population.

4 Experiments in cell membrane characterization

The EPR center of the Department of Solid State Physics at the Jožef Stefan Institute explores the lateral heterogeneity in various biologically and physiologically important systems. Examples of the inspected systems include human and other

mammalian blood cells, nerve cells and muscle cells, as well as other eucariotic and procariotic cells exposed to or modified by biochemical and biophysical stress conditions, such as various diseases and other environmental changes. Much of the research and experimental work is done in cell membrane characterization since the knowledge on the membrane structure and functionality could substantially improve our understanding of these complex systems. We have tested the new optimization approach to EPR spectral parameter optimization on labeled with lipophylic spin-probes. The membranes mimic the nerve cell membranes and therefore represent an important example of lateral heterogeneity found in real biomembranes.

4.1 Experimental setup

For the cell membrane characterization tests we used liposomes prepared from phosphatidilcholine (we denote this sample as Lipo(PL)) and the mixture of phosphatidilcholine and gangliosides in molar ratio 4:1 (Lipo(PL+GL)). They were all prepared in phosphate buffer solution (PBS) at pH = 7.4. The model membranes were labeled with MeFASL (10,3) – spin labeled methyl ester of palmitic ester using thin film preparation with ethanol evaporation. The label-to-lipid ratio was less than 1:270. The measurements were performed in glass capillary at 9.6 GHz EPR spectrometer at 35°C. For an additional variant of the problem the model membranes were prepared in 30% sucrose PBS buffered solution (we refer to this instance as Lipo(PL+GL)–SUC). For each of the three test problems the EPR spectrum consisting of 1024 data points was recorded. In spectrum simulation, three-domain models were assumed where 17 parameters were subject to optimization.

The hybrid evolutionary algorithm was run with the population size of 200, number of generations 100, crossover probability 0.7, mutation probability 0.05, and hybridization operator probability, i.e. the execution of downhill simplex on randomly selected solutions, 0.01. The optimization algorithm explored the parameter search space shown in Table 1 for each spectral domain.

Two sets of tests were performed. Initially, the robustness of the optimization

Table 1: Search space considered in parameter optimization for each of the three spectral domains.

Spectral parameter	Unit	Min. value	Max. value	Step size
Order parameter S	–	0.02	1	0.005
Rotation correlation time τ_c	ns	0.1	3	0.05
Broadening constant W	mT	0.02	0.3	0.005
Polarity correction factor p_A	–	0.8	1.2	0.001
Polarity correction factor p_g	–	0.9998	1.0002	0.000002
Weighting factor d	–	0.01	0.99	0.005

450   *Simulations in Biomedicine V*

technique was verified on Lipo(PL). Afterwards, spectral parameter optimization was performed on all three problem instances and the results compared with a previous version of the parameter tuning software using the downhill simplex method.

4.2 Testing the robustness of the optimization technique

Tuning of the spectrum simulation model parameters actually means fitting empirical data with the model. A potential weakness of data fitting techniques is overfitting which manifests in producing highly accurate results on input data, but performing poorly on previously unseen data related to the same phenomenon. To test the robustness of our parameter optimization technique against overfitting, we used the cross-validation method known from the field of data mining [8]. The experimentally obtained spectral points were randomly divided into ten subsets. Each subset representing 10% of the entire data was held out in turn, the spectral parameters were optimized on the remaining 90% of the data, and the quality of fit was assessed on both samples according to the χ^2 measure given by eqn (1). A sequence of ten such numerical experiments was performed, representing ten-fold cross validation.

This test was performed on the Lipo(PL) cell membrane characterization problem. The results for individual folds are given in Table 2. The average χ^2 values are 9.3 for the training subsets and 9.7 for the test subsets. These results confirm that the fit found does not critically depend on the fact that 10% of the data points were not used in optimization. Similar χ^2 values for the training and test data subsets also indicate that the optimized spectral parameters do not overfit the empirical data.

Table 2: Values of χ^2 obtained in tenfold cross-validation on the cell membrane characterization problem Lipo(PL).

Data subset	1	2	3	4	5	6	7	8	9	10
Training (90%)	7.4	14.1	9.5	6.9	7.7	7.5	10.5	10.4	10.3	9.1
Test (10%)	7.9	10.5	10.4	9.2	9.1	7.3	11.5	7.8	11.2	11.9

4.3 Tuning spectral parameters

The hybrid evolutionary optimization technique was applied to the three liposome cell characterization problems. For its stochastic nature, the algorithm was run ten times on each problem and the results were assessed statistically. In addition, its performance was compared with a human-navigated local optimization technique. This was a stand-alone downhill simplex optimizer, routinely used for EPR spectral parameter optimization before the new technique was introduced. The results are summarized in Table 3. It can be seen that the evolutionary optimization method produces repeatable results and outperforms the previously used

optimization tool not only in the best case but also on average. Further analysis of the results showed that the actual parameter values found by the two approaches are comparable.

Table 3: Values of χ^2 obtained in cell membrane characterization on three variants of the liposome problem.

Optimization procedure	Lipo(PL)	Lipo(PL+GL)	Lipo(PL+GL)–SUC
Downhill simplex	8.6	46.9	16.55
Hybrid EA, average	8.4±0.9	40.2±5.8	15.1±2.2
Hybrid EA, best	7.7	10.5	12.1

The most important practical implication of the new optimization technique is, however, a substantial reduction of the time spent by a spectroscopist tuning spectral parameters. Traditional single-point optimization techniques, like the one used previously in our laboratory experiments, require the preparation of promising initial solutions and navigation of the search according to the empirical knowledge on the properties of the inspected system. This may sometimes take hours to fit a single EPR spectrum. On the other hand, the hybrid evolutionary procedure starts from randomly selected starting points and runs automatically. It needs no human assistance and therefore releases a spectroscopist from time-consuming optimization and allows him/her to focus on the experiment.

5 Conclusion

A hybrid evolutionary algorithm for optimization of parameters in EPR spectroscopy was presented along with the results demonstrating its robustness against overfitting and its performance in cell membrane characterization. It was evaluated on real EPR spectra of liposome membranes where tuning corresponding biophysical models required 17 spectral parameters to be optimized. An important advantage of the hybrid procedure is that it provides a means for automating spectral parameter optimization. Unlike with less powerful search techniques, the spectroscopist is no more required to preprocess the solutions and intervene during the optimization procedure.

Further improvements and studies of applicability of evolutionary algorithms in EPR spectroscopy are underway. We experiment with knowledge-based genetic operators accounting for partially known correlations among spectral parameters. Using these in the stochastic search procedure is expected to result in higher quality of the fits and/or faster convergence of the optimization procedure. The new technique is now in experimental use and undergoes tests on a variety of real spectra of various origins and complexities. These activities are aimed at refining the current test version into a robust tool to effectively support characterization of biological systems with EPR spectroscopy.

452   *Simulations in Biomedicine V*

Acknowledgement

References

[1] Mouritsen O. G. & Jørgensen K. Dynamical order and disorder in lipid bilayers. *Chemistry and Physics of Lipids*, **73**, pp. 3–25, 1994.

[2] Marsh, D. Lipid-protein interactions and heterogeneous lipid distribution in membranes. *Molecular Membrane Biology*, **12**, pp. 59–64, 1995.

[3] Bäck, T., Fogel, D. B. & Michalewicz, Z. (eds). *Handbook of Evolutionary Computing*, Institute of Physics Publishing: Bristol, Philadelphia, and Oxford University Press: New York, Oxford, 1997.

[4] Filipič, B. & Štrancar, J. Tuning EPR spectral parameters with a genetic algorithm. *Applied Soft Computing*, **1(1)**, pp. 83–90, 2001.

[5] Filipič, B. & Štrancar, J. A hybrid evolutionary algorithm to facilitate characterization of biological systems with electron paramagnetic resonance, eds. R. Matoušek & P. Ošmera, *Proceedings of the 7th International Conference on Soft Computing, Mendel 2001*, University of Technology, Faculty of Mechanical Engineering: Brno, Czech Republic, pp. 125–130, 2001.

[6] Štrancar, J., Šentjurc, M. & Schara, M. Fast and accurate characterization of biological membranes by EPR spectral simulations of nitroxides. *Journal of Magnetic Resonance*, **142(2)**, pp. 254–265, 2000.

[7] Nelder, J. A. & Mead, R. A simplex method for function minimization. *Computer Journal*, **7(4)**, pp. 308–313, 1965.

[8] Witten, J. & Frank, E. *Data Mining, Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann: San Francisco, 2000.