Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 189 (2007) 139–150
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Deducing 1D concentration profiles from EPR imaging: A new approach based on the concept of virtual components and optimization with the genetic algorithm

Tomasz Spałek [a], Krzysztof Kruczała [a], Zbigniew Sojka [a,b,*], Shulamith Schlick [c]

[a] Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland
[b] Regional Laboratory of Physicochemical Analyses and Structural Research, Ingardena 3, 30-060 Krakow, Poland
[c] Department of Chemistry and Biochemistry, University of Detroit Mercy, 4001 West McNichols, Detroit, MI 48221-3038, USA

Received 20 March 2007; revised 24 July 2007
Available online 24 August 2007

Abstract

Application of the genetic algorithm (GA) in conjunction with the concept of virtual components (VC) to determine 1D concentration profiles from EPRI spectra (images) is described. In this approach the concentration profile is expressed as the superposition of virtual components described by analytical functions of the Gaussian and Boltzmann type. The method was implemented in the computer program ***ACon***, which allows for fully automated profile extraction via the nonlinear least-squares fitting of experimental images. The parametric sensitivity of the GA internal parameters such as population size, probabilities of the crossover, mutation and elitist retention to the search space was investigated in detail in order to find their optimal settings. The customized genetic algorithm was evaluated using simulated and experimental test data sets and its performance was compared with the Monte Carlo approach.

Keywords: Genetic algorithm; EPRI; ESRI; Imaging; Optimization; Concentration profile; Fitting; Virtual components

1. Introduction

Electron Paramagnetic Resonance Imaging (EPRI) is a powerful non-invasive technique for the determination of the spatial distribution of spins in the investigated samples [1,2]. Until the early 1980s the EPR spectra were usually measured with tacit assumption that the spin distribution could be treated as approximately uniform. However, since the pioneering paper of Hoch and Day describing the repartition of color centers in natural diamonds [3], mapping of the spin distribution in vast range of materials became a well-established field of the EPRI applications. Major advances have been made in biology, polymer degradation research [4–7] and spin dosimetry [8–10]. For a more comprehensive account, monographs or specialized review articles can be recommended [5,11,12].

The 1D image (the EPRI spectrum) is a convolution of the gradient-off spectrum with the concentration profile along the sample depth. A common approach for the deconvolution procedure is based on the Fourier transform (FT), followed by point-by-point least-squares refinement of the provisional profile extracted by FT analysis using Monte Carlo method [6,13]. However, such procedure often gives rise to high noise, especially when sharp edges in the profile are present. To circumvent this problem, the concentration profile obtained by FT has been fitted by analytical functions and convoluted with the gradient-off EPR spectrum to get the best agreement with the experimental image [14]. As the number of parameters to be optimized in such a procedure may be large for some complicated multimodal concentration distributions, a robust and efficient fitting subroutine is an indispensable component of the EPRI software. The complexity of the problem,

* Corresponding author. Address: Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland. Fax: +48 12 634 05 15.
E-mail address: sojka@chemia.uj.edu.pl (Z. Sojka).

1090-7807/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2007.08.012

however, imposes several restrictions on suitable optimization methods for profile determination, which should avoid computationally time-consuming derivative information, be resistant to local minima, and be effective in nonlinear problems. These requirements disfavor direct search methods that usually lead to suboptimal solutions, in favor of global techniques such as simulated annealing [15] or genetic algorithms (GA) [16,17] that tackle difficult optimization problems.

The genetic algorithm is an optimization method employing a probabilistic, non-local search heuristics that patterns natural evolution. The major strength of GA is the ability to explore and exploit a large poorly known parameter space with no initial guesses and derivative information, without trapping in local minima. As a result, genetic algorithms are of growing interest in various fields of science [18–20], with many successful applications in spectroscopy [21–23]. Recent notable examples are provided, for instance, by Meerts et al. [24–26]. There is also a growing number of GA applications in the field of EPR spectroscopy, including optimization of static multicomponent powder spectra [27–29] and motionally-restricted fluid-solution spectra [30–32].

In this study, we present a new VC-GA approach to automate the analysis of EPRI spectra. To eliminate the deconvolution noise, the 1D concentration profiles, expanded in terms of virtual components (VC) described by Gaussian functions, are convoluted with the gradient-off spectrum, and fitted to the experimental image using the genetic algorithm optimization. To our knowledge, the only successful application of GA for the determination of 1D concentration profiles was reported in our studies on the HPEC polymer degradation [33].

2. Theory

2.1. Determination of the concentration profile

As mentioned above, two spectra are needed to determine the 1D concentration profile from EPR imaging: the regular EPR spectrum and the 1D image (or EPRI spectrum) measured in the presence of the magnetic field gradient. The 1D image, $\boldsymbol{I}(\boldsymbol{B})$, is a convolution of the EPR spectrum $\boldsymbol{S}(\boldsymbol{B})$ registered in the absence of gradient with the function $\boldsymbol{Y}(\boldsymbol{r})$ describing spatial distribution of paramagnetic centers (the profile) along the gradient direction $\boldsymbol{r}$, modified by the longitudinal sensitivity response of the cavity, $\boldsymbol{C}(\boldsymbol{r})$:

$$\boldsymbol{I}(\boldsymbol{B}) = \int_{-\infty}^{+\infty} (\boldsymbol{S}(B - r \cdot \mathrm{grad}_r B) \cdot \boldsymbol{Y}(\boldsymbol{r}) \cdot \boldsymbol{C}(\boldsymbol{r}))\,\mathrm{d}r \qquad (1)$$

The convolution is legitimate (i.e., gives the right image) if the gradient-off EPR spectrum exhibits no spatial dependence. For short samples (less then 4 mm in X band), the sensitivity response of the resonator can be treated as being practically uniform, $\boldsymbol{C}(\boldsymbol{r}) = \mathrm{const}$. Then in a shorter form this equation can be rewritten as $\boldsymbol{I} = \boldsymbol{S} * \boldsymbol{Y}$ where $*$ is the convolution symbol.

A convenient procedure for solving the deconvolution problem is based on the Fourier transform (FT) theorem, which states that the Fourier transform of the convolution of two functions is equal to the product of their Fourier transforms ($F(\boldsymbol{I}) = F(\boldsymbol{S})F(\boldsymbol{Y})$) [6]. The problem is not trivial from the mathematical point of view and, as already mentioned, this method results in high frequency noise in the calculated profile, inducing some artifacts in the abruptly changing regions as well.

The inverse problem may be reformulated as an optimization problem. For solving Eq. (1), the a priori known spectrum $\boldsymbol{S(B)}$ can be regarded as a kernel function used for extraction of the searched profile $\boldsymbol{Y(r)}$ from the experimental image $\boldsymbol{I(B)}$ registered at a given magnetic field gradient [34]. However, the resultant inverse problem does not fulfill the Hadamard postulate, and is inherently mathematically ill-posed [35]. The solutions are highly unstable with respect to even small perturbations in the input data, and special regularization techniques are needed to obtain sensible results. In the nonlinear case, the corresponding objective function may be highly complicated, exhibiting multiple local optima, and often making the minimization processes a nontrivial issue. The success of the traditional local optimization methods strongly depends on the starting point [36,37], and is rapidly restrained with increasing number of the parameters, often failing to locate the global minimum. Thus, a promising solution of such problems can be obtained using a robust search, like that provided by the genetic algorithm.

2.2. Optimization procedure

Mathematically the calculated concentration profile, $\boldsymbol{Y}$, can be represented by a set of discrete data points $\boldsymbol{Y} = (Y_1 \ldots Y_n)$. It can next be expanded in terms of virtual components (VC) described by a series of Gaussians supplemented by two boundary functions of the Boltzmann type that allow for more accurate reproduction of steep changes at the edges of the sample.

$$\boldsymbol{Y} = B(r_1, a_1, h_1) + \sum_{k=2}^{n-1} G(r_k, a_k, h_k) + B(r_n, a_n, h_n) \qquad (2)$$

with $G(r_k, a_k, h_k) = a_k \cdot \exp[-(r - r_k)^2/2h_k^2]$ and $B(r_k, a_k, h_k) = a_k/[1 + \exp(r - r_k)/h_k]$, where r_k, a_k and h_k describe the position, the contribution coefficient and the half-width of the virtual component k. Use of analytical functions for this purpose has been proved elsewhere [14]. For convenience, the extension of the profile can be rescaled into the $[0, 1]$ range.

Following Eq. (1), the overall calculated profile can be next convoluted with the kernel function (the gradient-off spectrum) to give the corresponding EPRI image $\boldsymbol{i}$. If

$\boldsymbol{P} = \{p_i^j\}$ represents a full set of the parameters of all virtual components needed to describe the overall profile, than $\boldsymbol{Y} = f(\boldsymbol{P})$. Thus, the profile extraction problem consists in finding such $\boldsymbol{P}$ to make the calculated image $\boldsymbol{i} = f(\boldsymbol{S}, \boldsymbol{P})$ as close as possible to the experiment. The quality of the fit can be characterized by the weighted sum of the square deviations (RMS) between the calculated ($\boldsymbol{i}$) and experimental ($\boldsymbol{I}$) images, normalized by the number of data points n. However, since the virtual components constituting a given profile have essentially no physical meaning, any explicit information about the specific values of the associated parameters $\boldsymbol{P}$ is actually not required. Therefore, for the sake of the convenience, they can be treated as latent (internal) tunable variables of the corresponding profile $\boldsymbol{Y}$, which are initialized and optimized automatically but not reported.

2.3. Adapting the genetic algorithm to automated analysis of EPRI spectra

Many variants of evolutionary algorithms have been developed [38–40]. Below we present a brief description on how to tailor the genetic algorithm described elsewhere [41], and to automate the process of determining 1D concentration profiles from EPRI spectra *via* computer least-squares fitting, based on the concept of virtual components (VC) expansion and genetic algorithm (GA) optimization (the VC-GA approach).

The structure of the algorithm is outlined in the flowchart shown in Fig. 1, while the notation is explained in Table 1. To tailor the GA machinery towards a particular application, one should provide the size m of the initial population $\Omega^t(m)$ of individuals G_i, $i = 1, \ldots, m$, the coding

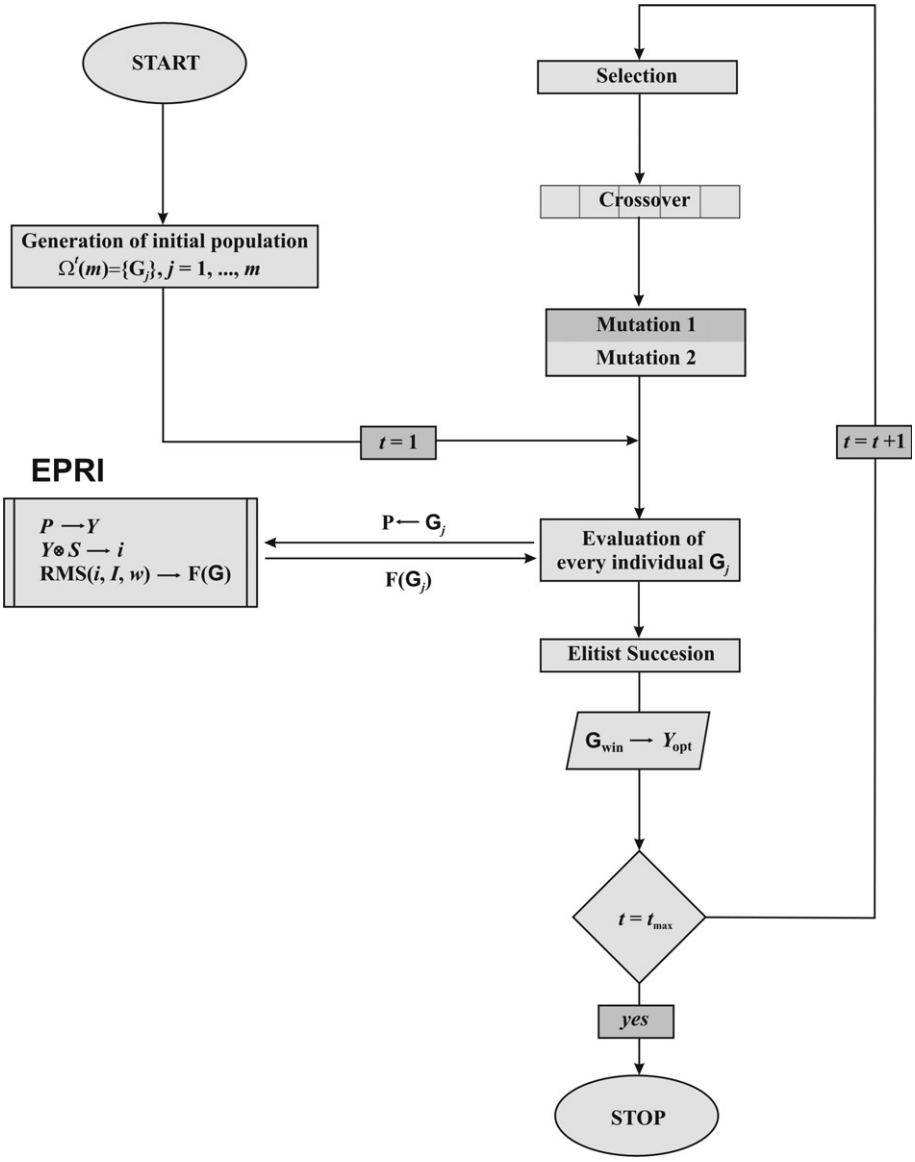

Fig. 1. Flowchart of the VC-GA method for automatic extraction of the 1D concentration profile from the EPRI spectra.

Table 1
Notations used in the flowchart of GA-based optimization of EPRI spectra

Notation	Definition
$\boldsymbol{I}$, $\boldsymbol{i}$	experimental and calculated EPRI spectra (images)
$\boldsymbol{Y}$	concentration profile
$\mathbf{S}$	experimental gradient-off EPR spectrum
$\boldsymbol{Y}$	concentration profile
$\boldsymbol{P} = \{p_i\}$	full set adjustable latent parameters p_i
$\Omega^t(m)$	tth population of the size m of individuals G_i
$\mathsf{G}_i = \{\mathsf{X}_1, \mathsf{X}_2, \ldots, \mathsf{X}_n\}$	an individual i composed of n chromosomes X_j corresponding to n virtual components constituting given concentration profile
$\mathsf{X}_j = [x_1, x_2, x_3] x_i \in [-1, 1]$	a string of chromosome with three genes corresponding to three parameters to be optimized
$F(\mathsf{G}_i)$	fitness index of an ith individual
$\mathrm{RMS}(\boldsymbol{i}, \boldsymbol{I}, \boldsymbol{w})$	root mean square error function
$\mathsf{G}_{\mathrm{win}}$	the fittest individual in the population t
G'_i	an individual i complemented by its fitness index $F(\mathsf{G}_i)$
$t_{\max}$	maximum number of generations

method, the fitness index $F(\mathsf{G}_i)$ of each individual, and the formulation of genetic operators.

In the adopted representation, a full set of the latent variables $\boldsymbol{P}$ to be optimized was encoded into an individual $\mathsf{G} = \{\mathsf{X}_1, \mathsf{X}_2, \ldots, \mathsf{X}_n\}$, composed of n chromosomes X_j corresponding to n virtual components of the overall profile. The polychromosomic individuals were preferred over the usual monochromosomic species, since they reflect better the structure of the optimization subject (reconstruction of the profile by superposition of virtual components). Each chromosome is a string of the form $\mathsf{X}_j = \{x_1^j, x_2^j, x_3^j\}$, where x_i^j stands for the genes representing the latent parameters p_i^j of the virtual component j that are optimized, but not reported overtly, as already mentioned. In the applied representation, the genes are coded as 32-bit floating point numbers from the $[0, 1]$ range. Although there is a lack of well established theoretical justification for using real instead of binary coded genes, most floating point optimizations seem to prefer real-valued versions of GA for solving high precision optimization problems when the parameters to be adjusted are continuous [42,43]. These real-valued versions of GA not only retain the proximity between two points in the representation and the problem spaces [41], but are more natural to implement than their binary analogs [44].

The adjustable internal parameters, i.e., the position r^j, contribution a^j and the width h^j, of the virtual component j are related to the corresponding genes x_i^j in the following way:

$$p_i^j = a_i x_i^j + b_i \tag{3}$$

where a_i and b_i ($i = 1, 3$) are empirically determined constants. All parameters are initialized randomly at the beginning of the optimization process. In such an approach the values of the latent variables are assigned without any prior knowledge about the profile structure, making the GA optimization a fully unsupervised method. Only the number of the virtual components and the type of the analytical functions (Gaussian or Boltzmann) have to be selected from the menu. The overall structure of such a representation is shown below in more detail.

Following the analogy with the gene–parameter relationship, the information contained in the chromosome X_j can be translated into the parameter space using the relation $P_j = \mathrm{A} \cdot \mathsf{X}_j + B$, where $P_j = \{p_1^j, p_2^j p_3^j)$, and $A = \{a_1, a_2, a_3\}$ and $B = \{b_1, b_2, b_3\}$ are the strings of the corresponding empirical constants. Consequently, the complete set $\boldsymbol{P}$ of latent parameters describing the profile $\boldsymbol{Y}$ is obtained by simple aggregation of the internal parameters of all partaking virtual components j, $\boldsymbol{P} = \{P_j\}, j = 1, \ldots, n$.

In this approach the ensemble of the chromosomes $\{\mathsf{X}_j\}$ constituting an individual G_i can be identified with a genotype encoding all profile parameters $\mathbf{G}_i \to \boldsymbol{P}$. The associated phenotype is the EPRI image derived from these parameters. Both spaces are linked together by the translation function (the convolution) converting the genotype space into the phenotype space ($\boldsymbol{i} = f(\boldsymbol{S}, \boldsymbol{P}(\mathbf{G}_i))$, which is in turn compared with the environment (the experimental image $\boldsymbol{I}$). The error function $\mathrm{RMS}(\boldsymbol{i}, \boldsymbol{I}, \boldsymbol{w})$ is used as a fitness criterion to be minimized (Fig. 2).

The flowchart shown in Fig. 1 explains the interplay between the GA optimization and the EPRI convolution program. At the beginning, an initial population $\Omega^1(m)$ of individuals G_i (potential solutions) has to be created randomly. Each gene of an individual G_i encodes a preliminary value of the associated adjustable parameter. The individuals are next evaluated by calling the program EPRI, and the calculated fitness values, $F(\mathbf{G}_i)$, are associated with the G_i strings ($\mathsf{G}_i \to \mathsf{G}'_i$). This is computationally the most time consuming step of the GA optimization, since it requires numerical simulation of the EPRI image. Actually the genetic algorithm makes use of a modified fitness function $F'(\mathsf{G}_i) = F(\mathsf{G}_o) - F(\mathbf{G}_i)$, where $F(\mathsf{G}_o)$ is the fitness of the worst adapted individual, i.e., that of the highest $\mathrm{RMS}(\boldsymbol{i}, \boldsymbol{I}, \boldsymbol{w})$ value in the initial population.

Based on the $F'(\mathsf{G}_i)$ values, the G'_i individuals are ranked according to their merit, and the fittest individual (the winner) is duplicated and saved as $\mathsf{G}_{\mathrm{win}}$. The set of parameters associated with $\mathsf{G}_{\mathrm{win}}$ is then used to calculate the overall concentration profile $\boldsymbol{Y}$, which is automatically displayed. Thus, at this stage the GA optimization is tantamount with the Monte Carlo search for an advantageous starting point, with the number of trials equal to the population size.

In the next step, the population is subjected to the reproduction process using the genetic operations in the sequence indicated in Fig. 1. The selection operation chooses the best individuals of the current population for reproduction in the next generation. A fitness proportionate selection (roulette wheel) [45] along with an elitist succession was applied to produce the progeny. In this

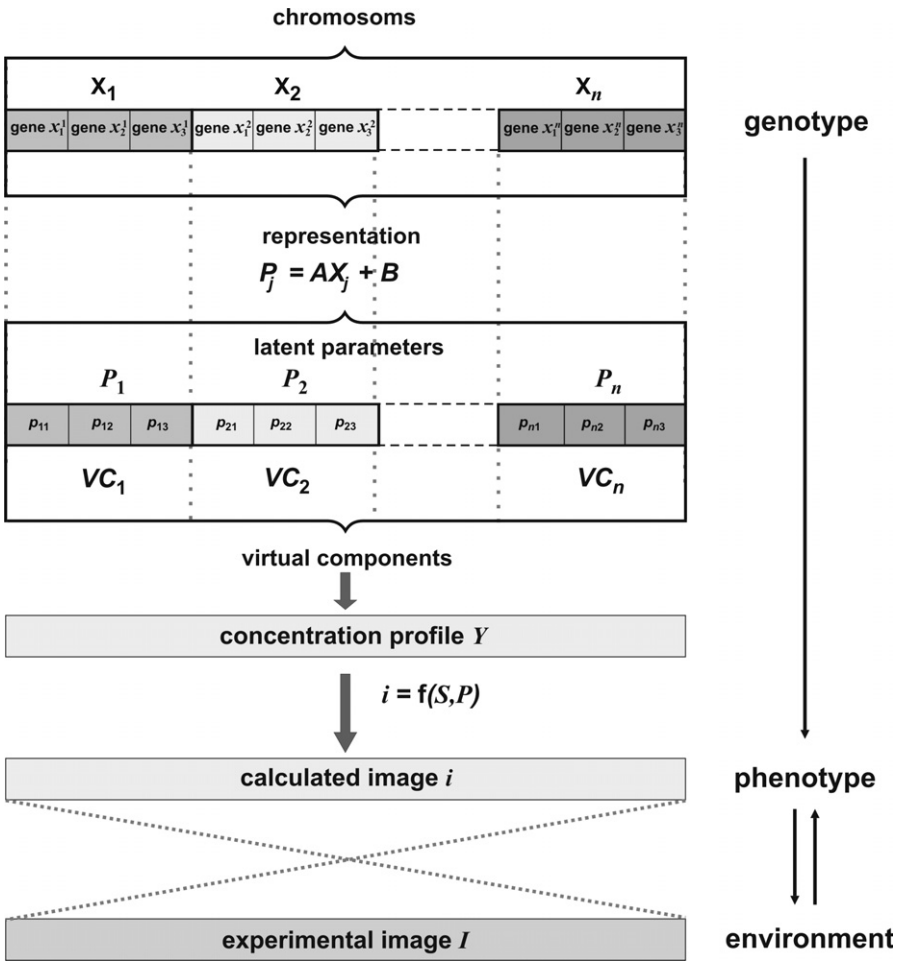

Fig. 2. The relationship between the genotype and phenotype spaces. The polychromosomic genotype of an individual G_i is mapped into the corresponding parameter space $\boldsymbol{P} = \{P_j\}$, defining the virtual components, VC, constituting the profile $\boldsymbol{Y}$. Next $\boldsymbol{Y}$ is used for convolution with the kernel spectrum $\boldsymbol{S}$ to give the EPRI image $\boldsymbol{i}$ (the phenotype) fitted to the experimental image $\boldsymbol{I}$ (the environment).

approach the reproduction probability $W_t(G_i)$ of an individual G_i is defined by a roulette wheel slot in proportion to its $F'(G_i)$ value, normalized over the whole population

$$W_t(G_i) = \frac{F'(G_i)_t}{\sum_{i=1}^{n} F'(G_i)_t} \tag{4}$$

To ensure that good solutions are not lost regularly, a fraction p_{elt} of the top individuals from the parent population is placed in the next generation (from which an equivalent number of the worst individuals were removed) without any changes. The actual fraction of retention defines the selection pressure. When the size of the elite is large from one generation to another, the algorithm can exhibit narrow population diversity. On the other hand, a small fraction might not retain enough of the desirable characteristics to overcome convergence problems. Therefore, the optimal p_{elt} value should be determined by appropriate meta-optimization of GA (vide infra).

The descendent individuals are produced through the operations of crossover and mutation. The recombination process was controlled by a probability criterion $p_{\text{crs}} > p_{\text{crs}}^{\text{thrs}}$, where $p_{\text{crs}}^{\text{thrs}}$ is an arbitrary threshold value for the crossover probability. A generalized k-point crossover operator was used for this purpose. The larger the number of the crossover points k, the more destructive is the corresponding operator. Five different crossover variants (a single point swapping and arithmetic, a two point swapping and arithmetic and a uniform crossover) were implemented in the GA [41]. The actual crossover option was selected at random at each run of the algorithm and the frequency of its use was determined by preset probabilities.

In the last step the chromosomes undergo the mutation operation. This operation is helpful in preventing premature convergence. A probability variable controls the mutation process, and usually has a small value inversely proportional to the number of genes [46]. We applied a sequential two-step mutation scheme invoking uniform and non-uniform gene variation, described by the following fluctuation function,

$$\Delta(\tau, z) = z\left(1 - r^{(1-\tau/t_{\max})^b}\right) \tag{5}$$

where τ stands for the evolution time, r is a random number from the $[0,1]$ range, z limits the maximum variation of the given gene, whereas b describes the system non-unifor-

mity [41]. It is well-known that a high mutation level at the beginning of a genetic algorithm (small values of τ) favors convergence by enhancing the exploration of the search space. However, at later stages of the optimization process (high values of τ), once the vicinity of the optimum is attained, high mutation levels may be counter-productive. However, to avoid any premature convergence to local subminima, the uniform mutation was permanently refreshing the pool of genes while the population evolves.

The resulting population is sorted by fitness and the top m individuals are retained for the next generation $\Omega^{t+1}(m)$. The iterative cycles (generations) are repeated until the optimization or the stopping criteria ($t \leqslant t_{\max}$) are fulfilled.

3. Program for concentration profile determination

The program, called ***ACon***, was written in Microsoft visual C^{++} 6.0 using the MFC library. The program based on the flow chart shown in Fig. 1 runs under Microsoft Windows with full 32 coding. Apart from the genetic fitting of the EPRI spectra based on the virtual components approach (GA-VC), the Fourier transform followed by point-by-point Monte Carlo refinement method (FT-MC) is also available. The ***ACon*** program has a user friendly interface based on mouse-driven access to pull-down menus and dialog boxes for fast communication. Interactive graphics for monitoring the simulation process and visualization of data in conjunction with the experimental EPRI spectra, including zoom and other standard display options, is available. Gradient-off spectra and gradient-on images, as well as the corresponding 1D concentration profiles can be saved as ASCII files, or in digital form. Hardcopy is available using any output device supported by Microsoft Windows. The program is available upon request.

4. Data sets

The performance of the GA optimization subroutine was tested using synthetic and experimental data sets of increasing complexity, corresponding to typical radical concentration profiles observed during accelerated aging of polymers [13,14]. A gradient-off isotropic test spectrum $\mathbf{T}_0$, ($g_{\text{iso}} = 2.00, a_{\text{iso}}/g\beta = 1.5$ mT) (Fig. 3a), represents the nitroxide radicals (HAS-NO) formed in the commercially available stabilizer Tinuvin 770, during polymer degradation.

The spectrum $\mathbf{T}_0$ was next convoluted with two types of profiles a rectangular one (Fig. 3b, left) and a bimodal one (Fig. 3c, left), giving EPRI test spectra (images) $\mathbf{T}_1$ and $\mathbf{T}_2$ (Fig. 3b (right) and c (right)). The efficiency of the GA fitting was also tested using experimental data sets. Fig. 4 (upper left) shows the gradient-off ESR spectrum $\mathbf{E}_0$ of HAS-NO radicals produced during thermal degradation of heterophasic propylene–ethylene copolymers (HPEC). Two HPEC samples differing in ethylene content were obtained from Dow Chemical Company: HPEC1 ($M_n = 60{,}700$, $M_w = 227{,}000$) and HPEC2 ($M_n = 90{,}400$, $M_w = 428{,}000$). The ethylene content was 25 wt.% in HPEC1 and 10 wt.% in HPEC2, ±2% [23]. The Tinuvin 770 stabilizer was provided by Ciba Specialty Chemicals. More information regarding the HPEC system can be found elsewhere [33]. The EPRI images of the samples after thermal degradation at various conditions ($\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$, $\mathbf{E}_4$), are presented on the left side of Fig. 4, whereas the images of the UVA-irradiated samples ($\mathbf{E}_5$, $\mathbf{E}_6$, $\mathbf{E}_7$, $\mathbf{E}_8$) are given on the right side of Fig. 4.

All experimental data were collected at 340 K with a Bruker EMX spectrometer operating at 9.4 GHz with 100 kHz magnetic field modulation, equipped with an Acquisit 32 Bit WINEPR data system version 3.01, and an ER 4111 VT variable temperature unit.

5. Meta-optimization of the intrinsic GA parameters

The operational efficiency of a genetic algorithm is intimately controlled by the values of such intrinsic parameters as the population size, the crossover and mutation probabilities, the decay factor and the percentage of elitist retention, which control various aspects of the algorithm in a given implementation. An optimal setting for these parameters would significantly enhance the general performance of the GA across different target sets. However, there is no sound theoretical basis for a priori assignment of their values, although some general hints are available [16,40]. Thus, in order to determine such settings for our application, parametric sensitivity tests were performed. Because of the stochastic nature of GA, several independent trials with randomly varying starting populations were performed in all testing experiments, in order to obtain reliable results and minimize the influence of the initial population.

5.1. Optimization of the population size

In general this parameter should be determined according to the specificity of the problem, and usually it is necessary to try out a range of the m values. We have tested this issue in more detail using $\mathbf{T}_1$ and $\mathbf{T}_2$ data sets. The size of the population was systematically varied from 25 to 500 with the step increment of 25, while all other parameters were kept constant at $p_{\text{crs}} = 0.8$ and $p_{\text{mut}} = 0.02$. Generally, the RMS values averaged over 10 runs and measured after the constraint calculation time, exhibited a moderate dependence on the population size with broad minima appearing at $m = 150$ and 200 for the $\mathbf{T}_1$ and $\mathbf{T}_2$ targets, respectively. However, the results were quite scattered, similarly to those previously observed in the case of powder EPR spectra, and discussed elsewhere [28]. In fact, the selection of the initial population size appeared to be less critical than the selection of the crossover and mutation rates (vide infra), and for further investigations the m value fixed at 200 was used in both cases.

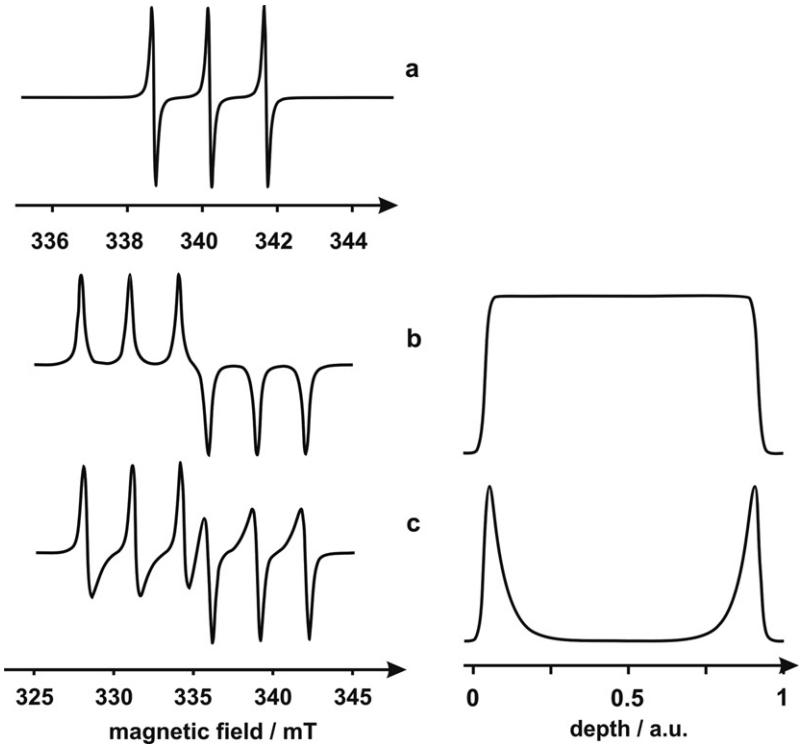

Fig. 3. Simulated test spectra: gradient-off EPR spectrum $\mathbf{T}_0$ (a), EPRI image (left) obtained from the convolution of $\mathbf{T}_0$ with a rectangular profile (right) (b), and with a bimodal test profile (right) (c).

5.2. Establishing the crossover and mutation rates

The sensitivity of the GA fitting of the EPRI images to the probabilities of crossover and mutation was examined using the $\mathbf{T}_1$ and $\mathbf{T}_2$ data sets. Optimization of these parameters is crucial because the relative rates of the crossover and mutation determine the balance between the exploitation and exploration moves, which are vital for the overall performance of the algorithm. The probability of the crossover (p_{crs}) was paced in the range $[0,1]$ with a 0.1 step, whereas the mutation probability (p_{mut}) was paced in the range $[0,0.1]$ with a 0.01 step. As a result, a grid of 121 data points was created and the GA optimization, with a population of size $m = 200$, was successively iterated over the generations for each pair of the p_{crs} and p_{mut} values. The results of the parametric sensitivity study, averaged over 10 independent runs, are summarized in the form of operator performance maps shown in Fig. 5, where the logarithm of the convergence efficiency (gauged by RMS) is drawn as a function of the crossover and mutation rates.

Inspection of Fig. 5 indicates that an increase of the crossover rate tends to enhance the performance of GA for the $\mathbf{T}_1$ and $\mathbf{T}_2$ targets. In the case of the test image $\mathbf{T}_1$, the optimal area is rather narrow and is located near $p_{\mathrm{crs}} \sim 0.8$. The optimal values for $\mathbf{T}_2$ are found in a broader range of the values (from 0.6 to 0.9), again with the minimum placed in the vicinity of $p_{\mathrm{crs}} = 0{,}8$. Crossover probabilities larger than 0.9 or smaller than 0.5 distinctly deteriorate the performance. The optimal range for the mutation probabilities is more confined in comparison to the crossover rate. Large ($p_{\mathrm{mut}} > 0.04$–0.03) and very small ($p_{\mathrm{mut}} < 0.01$) rates become harmful quite rapidly. For test spectra, the advantageous values appear at $p_{\mathrm{mut}} \sim 0.02$. The results of the sensitivity studies for the crossover and mutation operators revealed a strong dependence of the GA performance on the preset probability values, and to a less extent on the nature of the investigated target. The mutation rates around 0.02 and the crossover rates in the range of 0.8 appear to be a reasonable optimal setting for further investigations.

5.3. Optimization of the elite size

Elitism affects the selection pressure in the population. Its optimal setting may considerably accelerate the speed of the convergence. The sensitivity of the GA performance to the elite size was investigated using the same test data sets as above. The fraction of the elitist retention was gradually varied from $p_{\mathrm{elt}} = 0$ to 0.50 by increments of 0.05, keeping all other parameters constant. The results, averaged over 10 independent runs, are shown in Fig. 6. For both test profiles the curves exhibit quite distinct minima at $p_{\mathrm{elt}} \sim 0.10$. Thus, the performance of GA has gained much, once optimal fraction of the fittest solutions is shielded from the apparently deleterious effects of the subsequent crossover operations, while the population evolves. Interestingly, the optimal p_{elt} value seems to be not very sensitive to the nature of the investigated targets. Thus, in the following tests, the elite size was fixed at a value corresponding to $p_{\mathrm{elt}} = 0.10$.

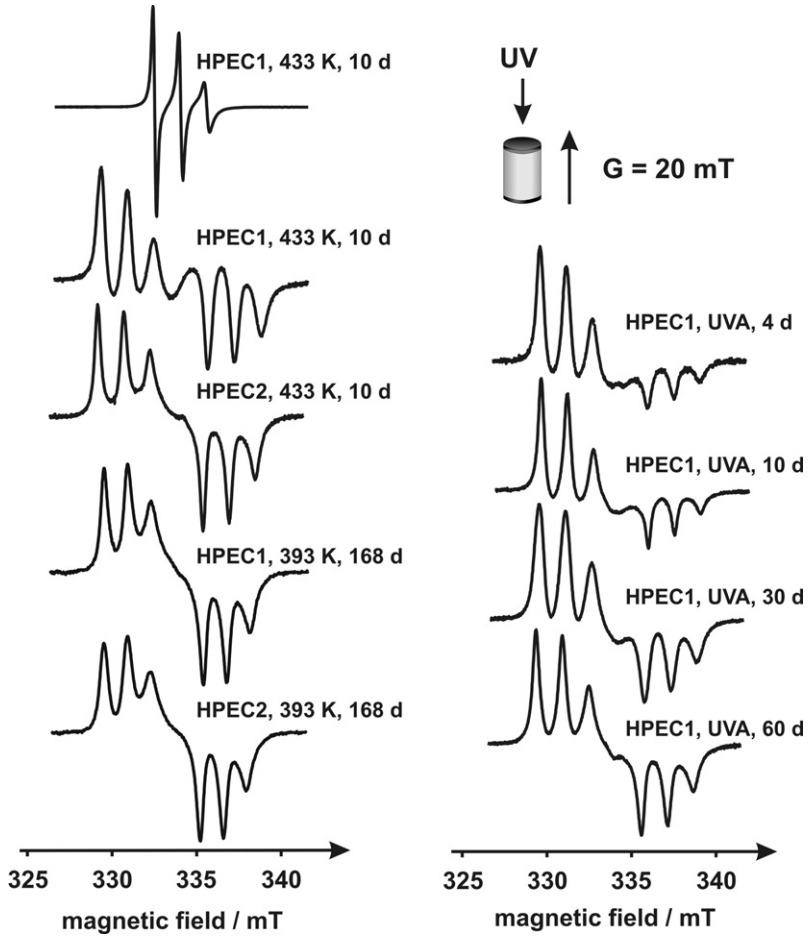

Fig. 4. Experimental EPR (left upper corner) and EPRI spectra registered at 340 K of thermally-treated at 433 K (left) and UVA-irradiated (right) HPEC polymers for the indicated conditions. The sample shape, the direction of UV irradiation, and of the magnetic field gradient (***G***) are shown in the upper right corner (experimental data taken from Ref. [33]).

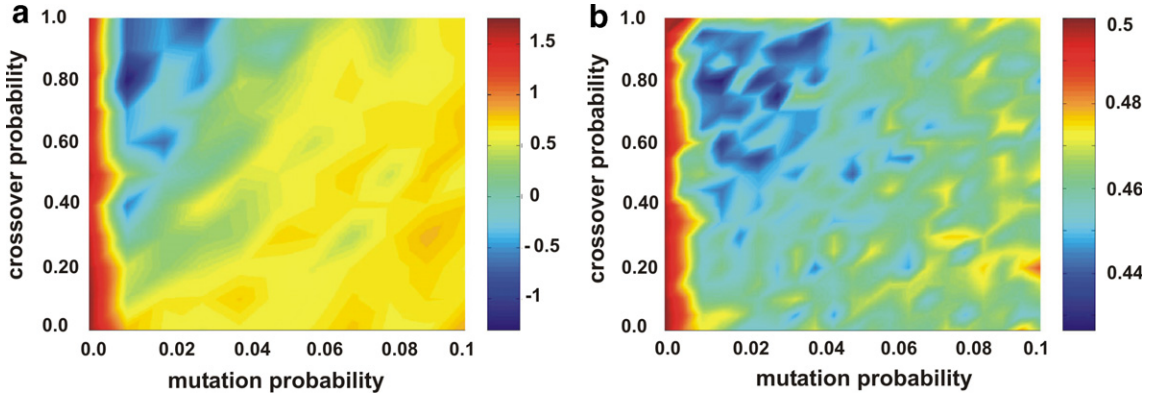

Fig. 5. Operator performance maps for parametric sensitivity of the GA convergence as a function of the mutation and crossover probabilities (a) for $\mathbf{T}_1$ and (b) for $\mathbf{T}_2$ targets. The color coding displays the convergence efficiency measured by the logarithm of the RMS value, averaged over 10 independent runs.

6. Determination of concentration profiles from 1D ESRI spectra

6.1. Simulated data

Undoubtedly, the first step in evaluating the performance of the GA optimization is to fit a simulated image derived from the a priori known profile. Only in such cases accurate solutions are achievable, which univocally identify the global minimum. Following this idea, we compared the performance of the GA optimization with the MC approach, using test profiles $\mathbf{T}_1$ and $\mathbf{T}_2$ convoluted with the $\mathbf{T}_0$ spectrum. The values of the GA parameters used in this study are listed in Table 2. For the convergence tests,

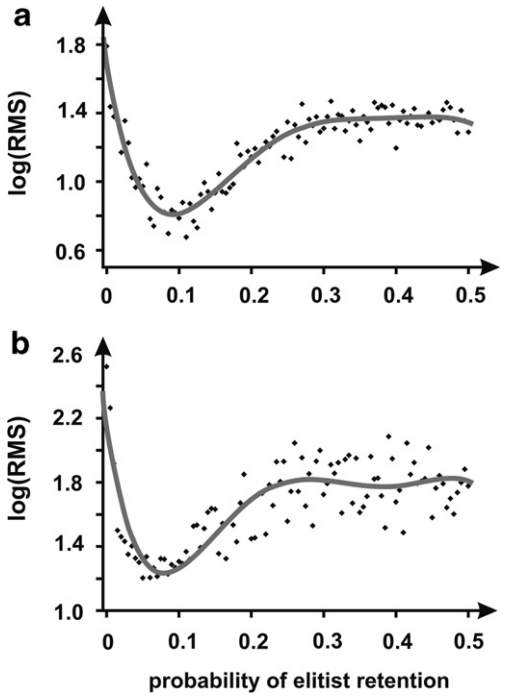

Fig. 6. Performance of GA as a function of the elitist retention for (a) $\mathbf{T}_1$ and (b) $\mathbf{T}_2$ test EPRI spectra.

Table 2
Settings of the genetic algorithm internal parameters

GA parameters	Setings
Population size	100
Crossover probability	0.8
Elitist retention	0.1
Mutation probability	0.02

both algorithms were initialized five times to equilibrate the results. The fitness evolution with the number of iterations for the $\mathbf{T}_0$ spectra convoluted with the rectangular and bimodal profiles respectively, is shown in Fig. 7a and b.

The green lines and red lines in Fig. 7 correspond to MC and GA minimization, respectively. In all cases the RMS values of the MC fitting were distinctly above those corresponding to the GA, and the optimization progress was quite irregular in nature. For the optimization with the genetic algorithm, decrease in RMS was exponential at the beginning, suggesting that the most rapid improvements occurred early in the evolutionary process. After $\sim 10^5$ iterations, the RMS values approaching the zero limit were reached in the case of the $\mathbf{T}_1$ target. For the more challenging $\mathbf{T}_2$ profile the performance of both algorithms was inferior, and the discrepancy between the progress of the MC and GA optimization dramatically amplified (Fig. 7b). After about 10^4 iterations of GA, the improvement of the fitting began to slow down, but still the values of RMS keep declining as the algorithm evolves. In Fig. 8a and b the resultant GA-calculated profiles (dotted lines) are compared with the test profiles (solid lines), showing an excellent agreement between the extracted and the test profiles in both cases.

Preliminary numerical experiments revealed that noise does not affect the quality of the recovered profiles appreciably, provided that signal-to-noise ratio is kept above 30 dB. Obviously, location of the global minimum in presence of considerable noise is a more difficult task, and may influence the convergence efficacy or lead to artificial profile distortions. More systematic investigations are needed to definitely resolve this issue.

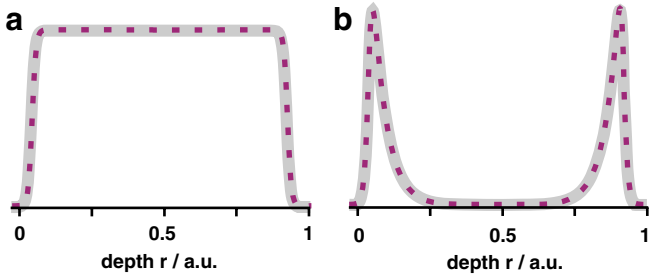

Fig. 8. Comparison of the original (gray solid line) and fitted (purple dotted line) profiles for (a) $\mathbf{T}_1$ and (b) $\mathbf{T}_2$ targets. (For interpretation of the references to color in this figure, the reader is referred to the web version of this paper.)

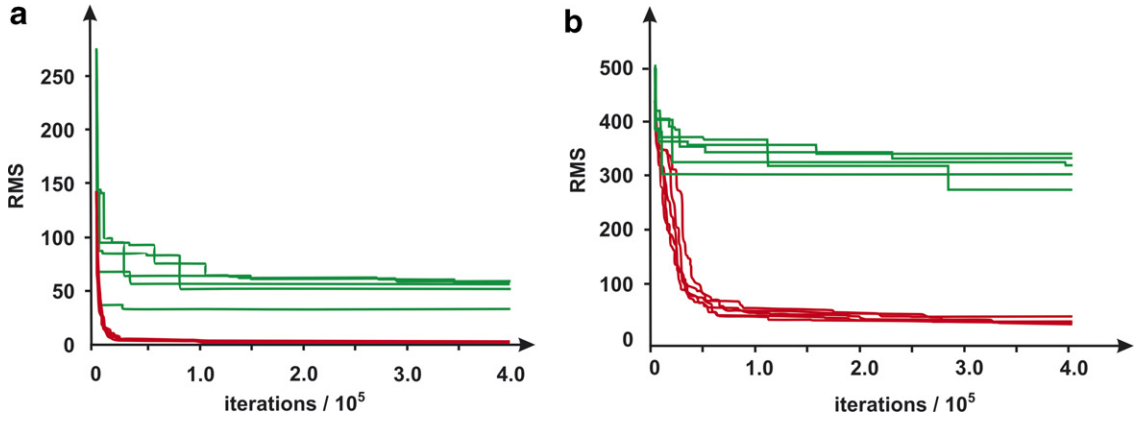

Fig. 7. Performance profiles for GA (red line) and MC (green line) optimization for (a) T1 and (b) T2 targets. (For interpretation of the references to color in this figure, the reader is referred to the web version of this paper.)

6.2. Experimental data

The GA optimization subroutine was next evaluated using the experimental data set $\mathbf{E}_1$–$\mathbf{E}_8$. The experimental and calculated images as well as the ensuing concentration profiles of the HAS-NO radicals produced in the HPEC samples upon thermal treatment or UVA-irradiation are shown in Figs. 9 and 10. In all cases the algorithm used in a totally automated way was able to precisely reproduce the experimental EPRI spectra, and to yield good quality concentration profiles as a result. This clearly demonstrates the capability and the power of the developed method. For thermally treated HPEC1 and HPEC2 samples, the HAS-NO radicals are almost homogenously distributed along the sample depth, producing a *quasi*-rectangular profile, except for the sample HPEC1 treated for 10 days at 433 K (Fig. 9a).

In the latter case a well-shaped bimodal concentration profile was obtained. The UV-degradation of the HPEC1 samples gave rise to a more complicated distribution of the radicals (Fig. 10), which was also sensitive to the exposure time. For the sake of the comparison between the VC-GA and FT-MC optimization, both extracted profiles are superimposed in Fig. 10c. As can be seen, their overall shapes are very similar; the clear advantage of the VC-GA approach is that it yields a completely noiseless profile, in contrast to FT-MC method. Undoubtedly, this benefit should be attributed to the VC ingredient of the ***ACon*** program, which allows for noise reduction due to its inherent smoothing ability. The robust GA optimization, however, assures that even in complicated cases, the concentration profiles can be extracted more reliably compared to MC, due to the better convergence.

7. Conclusions

This paper describes the implementation of a genetic algorithm for reconstruction of 1D concentration profiles via EPRI spectra fitting. Noiseless profiles can be recovered successfully by application of the concept of virtual components (VC). The efficacy of the method was demonstrated using synthetic and experimental data sets. The developed VC-GA approach succeeds in determining complex concentration profiles in a fully automated way without any prior knowledge about their shapes. The algorithm was implemented into the program ***ACon***, which, owing to

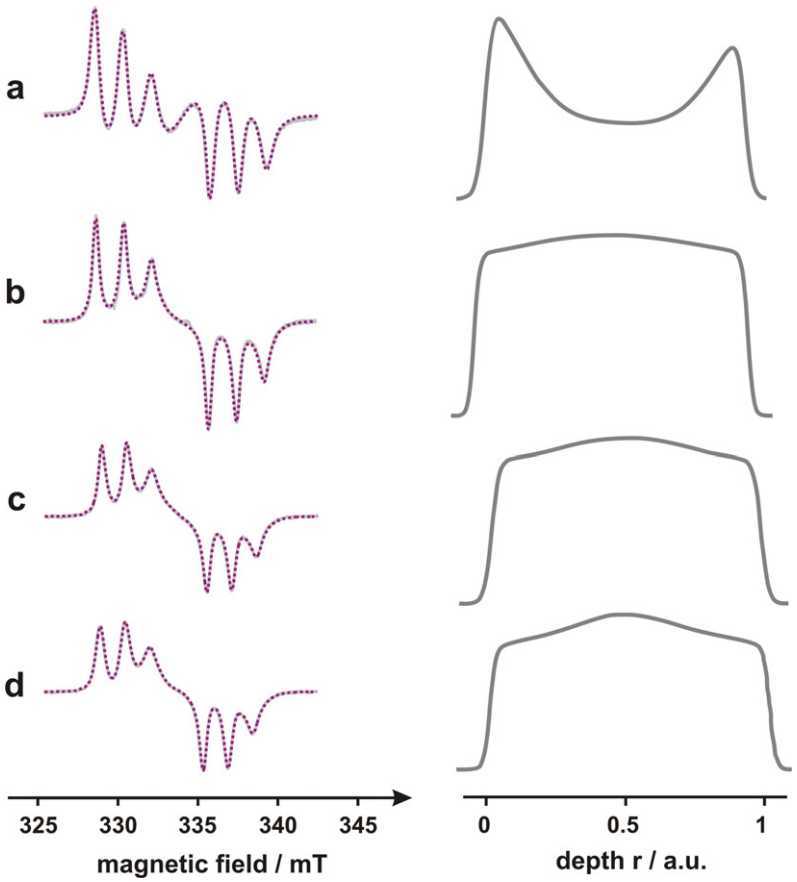

Fig. 9. Comparison of the experimental and simulated EPRI test spectra of thermally degraded HPEC polymer registered at 340 K. (a) EPRI spectrum of HPEC1 sample after 10 days of degradation at 433 K. (b) EPRI spectrum of HPEC2 sample after 10 days of degradation at 433 K. (c) EPRI spectrum of HPEC1 sample after 168 days of degradation at 393 K. (d) EPRI spectrum of HPEC2 sample after 168 days of degradation at 393 K. The corresponding concentrations profiles are given on the right side (experimental data taken from Ref. [33]).

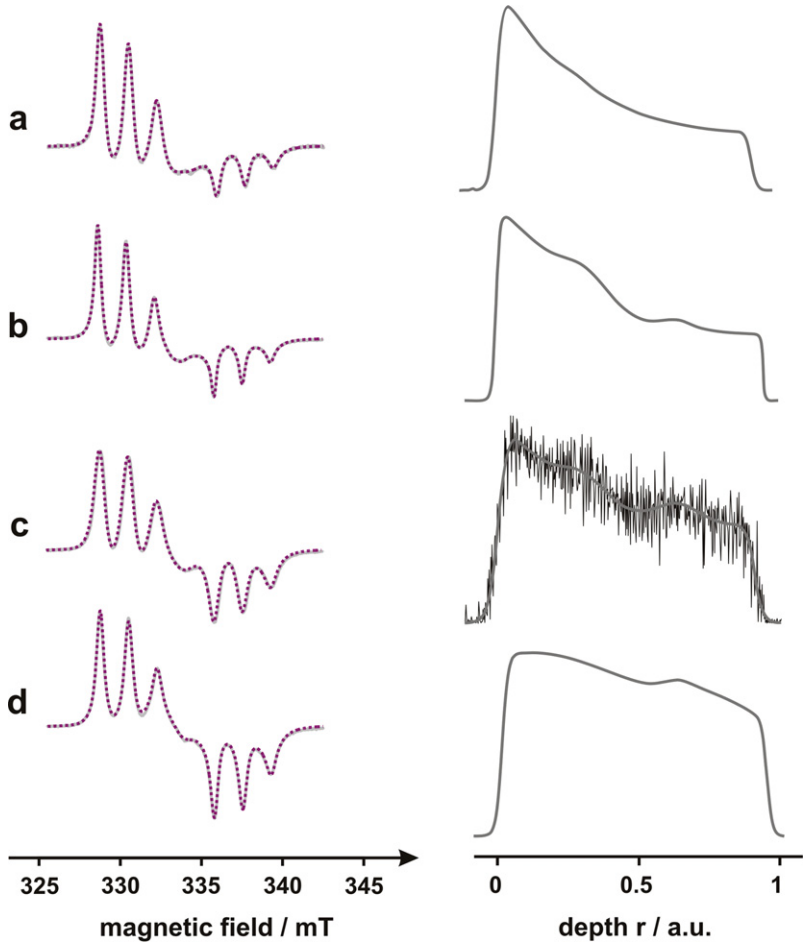

Fig. 10. Comparison of the experimental and simulated EPRI spectra of UV-degraded HPEC1 polymer registered at 340 K after (a) 4 days of irradiation, (b) 10 days of irradiation, (c) 30 days of irradiation and (d) 60 days of irradiation. The corresponding concentrations profiles are given on the right side (experimental data taken from Ref. [33]).

the user-friendly interface, allows fast and convenient analysis of EPRI data. The promising results reported herein speak in favor of further development of the VC-GA optimization to improve its convergence and elucidate the effect of the signal noise.

Acknowledgments

This study was supported in Poland by the Ministry of Education and Science of the Republic of Poland, Grant No. 3 T09A 051 28 and in USA by The Polymers Program of the National Science Foundation.

References

[1] A. Blank, C.R. Dunnam, P.P. Borbat, J.H. Freed, High resolution electron spin resonance microscopy, J. Magn. Reson. 165 (2003) 116–127.

[2] M. Ikeya, Electron-Spin-Resonance (ESR) microscopy in materials science, Annu. Rev. Matter. Sci. 21 (1991) 45–63.

[3] M.J.R. Hoch, A.R. Day, Imaging of paramagnetic centers in diamond, Solid State Commun. 30 (1979) 211–213.

[4] M.K. Ahn, S.S. Eaton, G.R. Eaton, M.A.B. Meador, Electron paramagnetic resonance imaging of the spatial distribution of free radicals in PMR-15 polyimide resins, Macromolecules 30 (1997) 8318–8321.

[5] S. Schlick, P. Eagle, K. Kruczala, J. Pilar, in: P. Blümler, B. Blümich, R. Botto, E. Fukushima (Eds.), Spatially Resolved Magnetic Resonance: Methods, Materials, Medicine, Biology, Rheology, Ecology, Hardware, Wiley, Weinheim, 1998, pp. 221–234 (Chapter 17).

[6] (a) M. Lucarini, G.F. Pedulli, V. Borzatta, N. Lelli, X-band EPR imaging of polymers irradiated with UV light, Res. Chem. Intermed. 22 (1996) 581;
(b) P. Franchi, M. Lucarini, G.F. Pedulli, M. Bonora, M. Vitali, Time evolution of the concentration profiles of HALS stabilizers and of the corresponding oxidation forms across polypropylene plaques irradiated with UV–visible light, Macromol. Chem. Phys. 202 (2001) 1246–1257.

[7] (a) M.V. Motyakin, J.L. Gerlock, S. Schlick, Electron spin resonance imaging (ESRI) of degradation and stabilization processes: behavior of a hindered amine stabilizer in UV-exposed poly(acrylonitrile–butadiene–styrene) (ABS) polymers, Macromolecules 32 (1999) 5463;
(b) M.V. Motyakin, S. Schlick, Electron spin resonance imaging (ESRI) and ATR–FTIR study of poly(acrylonitrile–butadiene–styrene) (ABS) containing a hindered amine stabilizer and thermally treated at 353 K, Macromolecules 35 (2002) 3984.

[8] D.A. Schauer, M.F. Desrosiers, P. Kuppusamy, J.L. Zweier, Radiation dosimetry of an accidental overexposure using EPR spectrometry and imaging of human bone, Appl. Radiat. Isotopes 47 (1996) 1345–1350.

[9] S.S. Ishchenko, S.M. Okulov, I.P. Vorona, Spatial distribution of radiation defects in tooth enamel, Phys. Solid State 41 (1999) 1100–1101.
[10] G.A. Watt, M.E. Newton, J.M. Baker, EPR and optical imaging of the growth-sector dependence of radiation-damage defect production in synthetic diamond, Diamond Relat. Mater. 10 (2001) 1681–1683.
[11] S.S. Eaton, G.R. Eaton, K. Ohno (Eds.), EPR Imaging and In Vivo EPR, CRC Press, Boca Raton, FL, 1991.
[12] Z. Xiang, Y. Xu, The status quo and prospect of ESR imaging applications to study on catalysts, Appl. Magn. Reson. 12 (1997) 69–79.
[13] K. Kruczala, M.V. Motyakin, S. Schlick, 1D and 2D spatial–spectral electron spin resonance imaging (ESRI) of degradation and stabilization processes in poly(acrylonitrile–butadiene–styrene) (ABS), J. Phys. Chem. B 104 (2000) 3387–3392.
[14] M.V. Motyakin, S. Schlick, Thermal degradation at 393 K of poly(acrylonitrile–butadiene–styrene) containing a hindered amine stabilizer: a study by 1D and 2D electron spin resonance imaging (ESRI) and ATR–FTIR, Polym. Degrad. Stab. 76 (2002) 25–36.
[15] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671–680.
[16] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, New York, 1989.
[17] L. Davis, Handbook of Genetic Algorithms, Van Norstrand Reinhold, New York, 1991.
[18] C.B. Lucasius, G. Kateman, Understanding and using genetic algorithms. Part 1: Concepts, properties and context, Chemometr. Intell. Lab. 19 (1993) 1–33.
[19] D.B. Hibbert, Genetic algorithms in chemistry, Chemometr. Intell. Lab. 19 (1993) 277–293.
[20] M. Moberg, D. Bylund, R. Danielsson, K. Markides, Optimization strategy for liquid chromatography–electrospray ionization mass spectrometry methods, Analyst 125 (2000) 1970–1976.
[21] J. Meiler, M. Will, Automated structure elucidation of organic molecules from ^{13}C NMR spectra using genetic algorithms and neural networks, J. Chem. Inf. Comput. Sci. 41 (2001) 1535–1546.
[22] M. Yang, X. Zhang, X. Li, X. Wu, A hybrid genetic algorithm for the fitting of models to electrochemical impedance data, J. Electroanal. Chem. 519 (2002) 1–8.
[23] B.M. Smith, P.J. Gemperline, Wavelength selection and optimization of pattern recognition methods using the genetic algorithm, Anal. Chim. Acta 423 (2000) 167–177.
[24] W.L. Meerts, M. Schmitt, G.C. Groenenboom, New applications of the genetic algorithm for the interpretation of high-resolution spectra, Canadian J. Chem. 82 (2004) 804–819.
[25] W.L. Meerts, M. Schmit, A new automated assign and analysing method for high-resolution rotationally resolved spectra using genetic algorithms, Phys. Scripta 73 (2006) C47–C52.
[26] W.L. Meerts, M. Schmitt, Application of genetic algorithms in automated assignments of high-resolution spectra, Int. Rev. Phys. Chem. 25 (3) (2006) 353–406.
[27] A. Adamski, T. Spałek, Z. Sojka, Application of EPR spectroscopy for elucidation of vanadium speciation in VO_x/ZrO_2 catalysts subject to redox treatment, Res. Chem. Intermediat. 29 (2003) 793–804.
[28] T. Spalek, P. Pietrzyk, Z. Sojka, Application of the genetic algorithm joint with the Powell method to nonlinear least-squares fitting of powder EPR spectra, J. Chem. Inform. Model. 45 (2005) 18–29.
[29] T. Spalek, P. Pietrzyk, Z. Sojka, Acta Phys. Pol. A 108 (2005) 95–102.
[30] B. Filipic, J. Strancar, Tuning EPR spectral parameters with a genetic algorithm, Appl. Soft Comput. 1 (2001) 83–90.
[31] J. Strancar, T. Koklic, Z. Arsov, B. Filipic, D. Stopar, M.A. Hemminga, Spin label EPR-based characterization of biosystem complexity, J. Chem. Inform. Model. 45 (2005) 394–406.
[32] A.A. Kavalenka, B. Filipic, M.A. Hemminga, J. Strancar, Speeding up a genetic algorithm for EPR-based spin label characterization of biosystem complexity, J. Chem. Inform. Model. 45 (2005) 1628–1635.
[33] (a) K. Kruczala, B. Varghese, J.G. Bokria, S. Schlick, Thermal aging of heterophasic propylene–ethylene copolymers: morphological aspects, based on ESR, FTIR and DSC, Macromolecules 36 (2003) 1899–1908;
(b) K. Kruczala, B. Varghese, J.G. Bokria, S. Schlick, Thermal aging of heterophasic propylene–ethylene copolymers: spatial and temporal aspects of degradation, based on ESR, ESR Imaging, and FTIR, Macromolecules 36 (2003) 1909–1919;
(c) K. Kruczala, W. Aris, S. Schlick, Stabilization and early degradation stages in UVA- and UVB-irradiated heterophasic propylene–ethylene copolymers studied by electron spin resonance imaging (ESRI): effect of ethylene content and UV wavelength, Macromolecules 38 (2005) 6979–6987.
[34] A. Shrestyuk, Kernel functions in convolution surfaces: a comparative analysis, Visual Comput. 15 (1999) 171–182.
[35] J. Hadamard, Lectures on Cauchy Problem in Linear Partial Differential Equations, Yale University Press, New Heavens, USA, 1923.
[36] W.A. Bernhard, G.W. Fouse, Simulation of powder ESR spectra of organic free radicals, J. Magn. Res. 82 (1989) 156–161.
[37] M. Griffin, A. Muys, C. Noble, D. Wang, C. Eldershaw, K.E. Gates, K. Burrage, G.R. Hanson, XSOPHE, a computer simulation software suite for the analysis of the electron paramagnetic spectra, Mol. Phys. Rep. 26 (1999) 60–84.
[38] L. Fogel, A. Owens, M. Walsh, Artificial Intelligence through Simulated Evolution, John Wiley, New York, 1966.
[39] (a) I. Rechenberg, Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution, Frommann–Holzboorg, Stuttgart, 1973;
(b) H.-P. Schwefel, Numerical optimization of computer models, John Wiley and Sons, New York, 1981.
[40] J. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Michigan, 1975.
[41] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer, Berlin, 1992.
[42] N.S. Mera, L. Elliot, D.B. Ingham, A Multi-population Genetic Algorithm Approach for Solving Ill-posed Problems, Comput. Mech. 33 (2004) 245–262.
[43] K.P. Ferentinos, K.G. Arvanitis, N.J. Sigrimis, Global Opt. 23 (2002) 155–170.
[44] L. Davis, Handbook of Genetic Algorithm, Van Nostrand Reinhold, New York, 1991.
[45] A.K.D. Jong, An Analysis of the Behavior of a Class of Genetic Adaptive Systems, Ph.D. thesis, University of Michigan, Ann Arbor, 1975.
[46] W.Y. Choy, B.C. Sanctuary, Using genetic algorithms with a priori knowledge for quantitative NMR signal analysis, J. Chem. Inf. Comput. Sci. 38 (1998) 685–690.