THE JOURNAL OF PHYSICAL CHEMISTRY A

 VOLUME 110, NUMBER 48, DECEMBER 7, 2006

ARTICLES

Combined Genetic Algorithm and Multiple Linear Regression (GA-MLR) Optimizer: Application to Multi-exponential Fluorescence Decay Surface

Jacek J. Fisz*

Institute of Physics, Nicolaus Copernicus University, ul. Grudziadzka 5/7, PL 87-100 Toruñ, Poland

Received: June 26, 2006; In Final Form: September 21, 2006

The optimization approach based on the genetic algorithm (GA) combined with multiple linear regression (MLR) method, is discussed. The GA-MLR optimizer is designed for the nonlinear least-squares problems in which the model functions are linear combinations of nonlinear functions. GA optimizes the nonlinear parameters, and the linear parameters are calculated from MLR. GA-MLR is an *intuitive* optimization approach and it exploits all advantages of the genetic algorithm technique. This optimization method results from an appropriate combination of two well-known optimization methods. The MLR method is embedded in the GA optimizer and linear and nonlinear model parameters are optimized in parallel. The MLR method is the only one strictly mathematical "tool" involved in GA-MLR. The GA-MLR approach simplifies and accelerates considerably the optimization process because the linear parameters are not the fitted ones. Its properties are exemplified by the analysis of the kinetic biexponential fluorescence decay surface corresponding to a two-excited-state interconversion process. A short discussion of the variable projection (VP) algorithm, designed for the same class of the optimization problems, is presented. VP is a very advanced mathematical formalism that involves the methods of nonlinear functionals, algebra of linear projectors, and the formalism of Fréchet derivatives and pseudo-inverses. Additional explanatory comments are added on the application of recently introduced the GA-NR optimizer to simultaneous recovery of linear and weakly nonlinear parameters occurring in the same optimization problem together with nonlinear parameters. The GA-NR optimizer combines the GA method with the NR method, in which the minimum-value condition for the quadratic approximation to χ^2, obtained from the Taylor series expansion of χ^2, is recovered by means of the Newton−Raphson algorithm. The application of the GA-NR optimizer to model functions which are multi-linear combinations of nonlinear functions, is indicated. The VP algorithm does not distinguish the weakly nonlinear parameters from the nonlinear ones and it does not apply to the model functions which are multi-linear combinations of nonlinear functions.

1. Introduction

The simultaneous (global) analysis of multiple fluorescence decay traces[1] is well-known methodology in fluorescence spectroscopy. Many different algorithms and their numerical implementations designed for the global and target analyses have been offered and tested in the literature,[2−9] which have found many fantastic practical applications in different problems of the fluorescence spectroscopy of solutions and ordered molecular systems.[10,11]

In our recent two articles (see refs 12 and 13) we have discussed the application of the genetic-algorithms-based (GA) optimization approach to time-resolved polarized fluorescence

* Corresponding author. E-mail: jjfisz@phys.uni.torun.pl.

10.1021/jp063998e CCC: $33.50
Published on Web 11/10/2006

spectroscopy of ordered molecular media. Genetic algorithms[14−16] are the class of heuristic optimization methods that involve the basic principles of the evolutionary biology. The GA method iteratively modifies a set of randomly generated trial population of the complete sets of the model parameters by employing in the reproduction process two genetic operators, i.e., crossover and mutation. At each iteration the "fitter" individuals in the population (i.e., those which return better χ^2 values) are kept for further reproduction, and the less "fit" ones are replaced by new randomly generated individuals. This procedure is repeated iteratively until χ^2 reaches a predefined tolerance or the number of iterations (generations) reaches its predefined maximum.

In ref 12 we have demonstrated the comparative numerical studies of the GA and gradient expansion (GE) optimization methods on the basis of polarized fluorescence spectroscopy of microscopically ordered membrane vesicles and macroscopically ordered planar membranes, which have displayed several very important advantages of the GA optimizer. First, in contrast to the GE optimizer, in the GA method no initial guesses for the fitted model parameters are required and only the upper and lower limits for the model parameters have to be predefined. Second, the GA optimization method is insensitive to the local minima of the χ^2 surface; i.e., the GA optimization process is not "trapped" at the local minima of the χ^2 surface. This is a very inconvenient property of the GE method, and in such cases the GE optimization procedure must be run several times with different initial guesses for the fitted parameters (see the illustrative examples discussed in ref 12). A very important property of the GA optimizer is that it is applicable also to the optimization problems in which the model functions are not differentiable in the entire space of fitted model parameters, because in the GA method no derivatives over the model parameters are evaluated. This is in evident contrast to the GE method for which such model functions represent serious difficulty. Also, an important property of the GA optimizer is that it can be assumed as a very efficient way for obtaining the initial guesses for GE optimization (e.g., in the case when the χ^2 surface possess very many local minima and when this surface is very flat around its global minimum), leading the GE optimization to almost immediate convergence to the very probable coordinates of the global minimum of χ^2 function.

In ref 13 we have discussed the methods for reduction of the number of fitted nonlinear and linear model parameters appearing in the nonlinear model functions. Such methods may eliminate essential problems with the "inconvenient" nonlinear and linear fitted parameters, for which the prediction of initial guesses in the GE optimization or their most adequate upper and lower bounds in the GA optimization represents a serious difficulty. We have discussed the combination of the GA optimization method with (a) the first-order derivative (FOD)[17,18] method (which applies to linear parameters, solely) and (b) the Newton−Raphson (NR)[17] method (which applies to both linear and nonlinear parameters). In the NR method the minimum of quadratic approximation to χ^2 function (obtained from the Taylor series expansion) is recovered iteratively by means of the Newton−Raphson (NR) algorithm. According to the GA-NR and GA-FOD optimization procedures, the fitted model parameters are split into those that are fitted by the GA optimization routine and the parameters that are being calculated from, respectively, the NR and FOD method, embedded in the GA routine (i.e., GA is an outer optimizer, and FOD and NR are the inner ones).

Although the numerical tests of the GA-FOD optimizer, demonstarted in ref 13, were concentrated on the global analyses of the polarized fluorescence decays with different single amplitudes (they were the set of linear parameters in the set of globally analyzed decays), in section 3 of that reference we have mentioned the applicability of the GA-FOD method to multicomponent nonlinear model functions given by linear combinations of the nonlinear functions, which, for example, can be related to kinetic and polarized fluorescence decays (for solutions and ordered media) of the compounds undergoing excited-state processes. In such cases the linear model parameters are evaluated from the multiple linear regression (MLR) method, and the whole algorithm combines the GA and MLR optimizers, yelding the GA-MLR optimization method, in which MLR is embedded in GA.

It is essentially important to mention here the variable projection (VP) algorithm[21,22] (and its numerical implementation VAPRO[22]) designed for the nonlinear least-squares analysis of the model functions that are linear combinations of nonlinear functions, and in which the linear parameters are estimated from the linear least-squares methods.

This method was unknown to us when submiting the manuscript of ref 13. VP is a very advanced mathematical formalism. It involves the methods of nonlinear functionals, algebra of linear projectors, and the formalism of Fréchet derivatives and pseudo-inverses. The VP formalism is based on a series of theorems and lemmas, proven in ref 21, which concern Fréchet derivatives of projectors, residual vectors and pseudo-inverses, and which represent a very strong mathematical background of this method. An important requirement in the VP approach is that the nonlinear functions must be continuously (at least twice) differentiable with respect to nonlinear parameters.

In the VP algorithm, the nonlinear functional (which can be referred to as a χ^2 function of the linear and nonlinear model parameters) is projected into the modified (variable projection) functional of a more complicated form but depends solely on nonlinear parameters. This projection converts the optimization procedure into the two-step one, which consists of first optimizing the nonlinear parameters (contained in the modified functional) and then using their optimal values obtained for solving the least-squares problem for linear parameters, with the application of the Moore−Penrose generalized inversion method.[21,22] In VAPRO, the optimization of nonlinear parameters is based on the modified Levenberg−Marquardt algorithm.[22] Many examples of different excellent applications of VP and VAPRO (and their modifications) are discussed in ref 22.

In this article we discuss in a systematic way and exemplify the combined genetic algorithm and multiple linear regression (GA-MLR) optimizer designed for the analysis of nonlinear model functions which are linear combinations of nonlinear functions. The GA-MLR optimization method is an *intuitive* treatment to nonlinear least-squares problems in which the nonlinear and linear parameters separate. This optimizer takes all advantages of the genetic algorithms and the whole method results from an appropriate combination of two well-known optimization methods. The GA-MLR optimizer does not require any assumption on the differentiability of the nonlinear constituent functions. They may have arbitrary forms (i.e., they may by given explicitly or numerically), and they may be arbitrarily complicated functions of the nonlinear parameters. These advantages of the GA-MLR optimizer result from the fact that MLR is the only one strictly mathematical "tool" involved in this optimization algorithm and from the fact that the nonlinear parameters are fitted by GA optimizer that possess particularly useful properties.

In the GA-MLR optimization approach, the term "combined" has exact meaning because MLR is embedded in GA and both optimizers converge simultaneously. The linear and nonlinear

fitted parameters are estimated in parallel, i.e., at each step of the GA-MLR algorithm. As the temporarily "better" values of the nonlinear parameters (e.g., the decay times) are selected, they, together with the linear parameters (e.g., the amplitudes at multiexponential decay) estimated from the MLR routine, return a better temporal goodness of the fit.

It is clearly seen from the above discussion that the only similarity between the VP and GA-MLR approaches is that they both are designed for the same class of least-squares problems, in which the linear and nonlinear parameters separate. From the point of view of their constructions and from the point of view of mathematical methods and tools they involve, VP and GA-MLR are completely different approaches.

One of the simplest classes of the model functions that can be subjected to the GA-MLR optimization method is the multiexponential kinetic and polarized fluorescence or transient absorption decays and decay surfaces, in the case of solutions and ordered media.[10,11]

We here formulate a general description of the GA-MLR optimization approach, not related to any particular optimization problem, and afterward, we discuss its adaptation to the global analysis of synthetic fluorescence decay surface for the case of a two-excited-state interconversion processes in solution, which can be referred to intramolecular charge transfer (ICT), twisted intramolecular charge transfer (TICT), or intramolecular proton transfer (PT) processes. In this example the fluorescence decay surface is composed of 71 biexponential fluorescence decays parametrized by 144 decay parameters., i.e., 2 decay times and 142 amplitudes. In the GA-MLR optimization method both decay times are the only fitted parameters (optimized by GA), and all 142 amplitudes are just calculated from the MLR method embedded in GA routine. By discussing this example, we want to demonstrate the simplicity and very high efficiency of the GA-MLR optimization method. Furthermore, we want to show that the GA-MLR optimizer when used together with the efficient methods for the analysis of the kinetic fluorescence decays for the ICT, TICT and PT processes, outlined in ref 23, allows for almost immediate recovery of the emission bands and the values of the kinetic rate constants, assumed in the simulations.

To make the discussion presented in this article more complete, we add a few important explanatory comments to the GA-NR approach, outlined in ref 13, as the method for simultaneous optimization of linear and weakly nonlinear model parameters that occur in the same optimization problem together with the nonlinear parameters. GA-NR optimizer combines the outer GA optimizer with the inner one NR, in which the minimum-value condition for the quadratic approximation to χ^2, obtained from the Taylor series expansion of χ^2, is explored by means of the Newton−Raphson algorithm. The inner optimizer NR is designed for the problems in which the set of separated parameters lead to nonquadratic dependence of χ^2. It applies to linear parameters occurring in the model functions that are multilinear combinations of nonlinear functions, to weakly nonlinear parameters and to nonlinear ones. The VP method does not distinguish the weakly nonlinear parameters (e.g., the quasi-linear parameters) from the nonlinear ones, and hence, VP does not apply to the optimization problems in which one wants to distinguish between these two sets of the model parameters. Furthermore, the VP method does not apply to the model functions that are multilinear combinations of nonlinear functions.

2. Genetic Algorithm Combined with Multiple Linear Regression (GA-MLR)

Let us assume that (x_i, y_i) (where $i = 1, ..., N$) represent the set of experimentally recovered N data points and that $y(\mathbf{q},\mathbf{s},x)$, given by

$$y(\mathbf{q},\mathbf{s},x) = \sum_{k=1}^{M} a_k F_k(\mathbf{q},x) \tag{1}$$

is the model function, which is a linear combination of the set of M nonlinear basis functions $F_k(\mathbf{q},x)$ of variable x and which are paramatrized by vector $\mathbf{q}$ of nonlinear model parameters. The combination parameters a_k form vector $\mathbf{s} = [a_1, ..., a_M]$ of linear model parameters in (1). The "best" estimates (a vector $\mathbf{p}_m$) of all model parameters $\mathbf{p} = [\mathbf{q}, \mathbf{s}]$ in the model function (1) are the coordinates of global minimum of the reduced χ^2-$(\mathbf{q},\mathbf{s})$ function[18,19]

$$\chi^2(\mathbf{q},\mathbf{s}) = \frac{1}{\nu_f} \sum_{i=1}^{N} \frac{(y_i - y(\mathbf{q},\mathbf{s},x_i))^2}{\sigma_i^2} \tag{2}$$

in space $\mathbf{p}$. ν_f is the number of degrees of freedom and σ_i is the measurement error (standard deviation) of the ith data point.

In the GA-MLR optimization approach, the GA optimizer provides a trial population of vectors $\tilde{\mathbf{q}}$ (i.e., a trial population of the complete sets of the nonlinear parameters in $F_k(\mathbf{q},x)$ ($k = 1, ..., M$)). Each trial vector $\tilde{\mathbf{q}}$ returns the corresponding χ^2-$(\tilde{\mathbf{q}},\mathbf{s})$ function of vector $\mathbf{s}$ of the combination parameters a_k standing at the corresponding function $F_k(\tilde{\mathbf{q}},x)$, namely

$$\chi^2(\tilde{\mathbf{q}},\mathbf{s}) = \frac{1}{\nu_f} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (y_i - \sum_{k=1}^{M} a_k F_k(\tilde{\mathbf{q}},x))^2 \tag{3}$$

Formally, all a_k can be determined from the MLR method, i.e., from the minimum-value condition of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ with respect to $\mathbf{s}$, that is, from[17−19]

$$\frac{\partial \chi^2(\tilde{\mathbf{q}},\mathbf{s})}{\partial \mathbf{s}} = 0 \tag{4}$$

This condition leads to the following set of simultaneous equations (which are the so-called normal equations of the least-squares problem)

$$\sum_{k=1}^{M} \left(\sum_{i=1}^{N} \frac{F_j(\tilde{\mathbf{q}},x_i)\, F_k(\tilde{\mathbf{q}},x_i)}{\sigma_i^2} \right) a_k - \sum_{i=1}^{N} \frac{y_i F_j(\tilde{\mathbf{q}},x_i)}{\sigma_i^2} = 0 \tag{5}$$

where $j = 1, ..., M$, which is equivalent to the matrix equation

$$\mathbf{As} = \mathbf{c} \tag{6}$$

where $\mathbf{s}$ is a column vector of the amplitudes α_k and

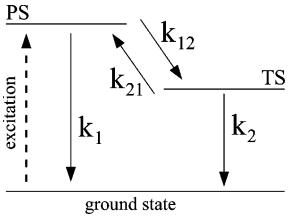
$$A_{jk} = \sum_{i=1}^{N} \frac{F_k(\tilde{\mathbf{q}},x_i)\, F_j(\tilde{\mathbf{q}},x_i)}{\sigma_i^2} \qquad c_j = \sum_{i=1}^{N} \frac{y_i F_j(\tilde{\mathbf{q}},x_i)}{\sigma_i^2} \tag{7}$$

A formal solution to eq 6 reads

$$\tilde{\mathbf{s}} = \mathbf{A}^{-1}\mathbf{c} \tag{8}$$

Substituting the obtained values of the amplitudes α_k into eq 3 and repeating this procedure over all trial vectors $\tilde{\mathbf{q}}$, the trial set of the values of $\chi^2(\tilde{\mathbf{q}},\tilde{\mathbf{s}})$ is obtained. As temporarily the "best"

SCHEME 1



vectors, $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{s}}$ in the population are considered to be those that correspond to the lowest χ^2 value in the set of all $\chi^2(\tilde{\mathbf{q}},\tilde{\mathbf{s}})$ obtained. The GA optimizer keeps for further reproduction those vectors $\tilde{\mathbf{q}}$ in the population that, together with the corresponding vectors $\tilde{\mathbf{s}}$ (estimated from the MLR method), return better $\chi^2(\tilde{\mathbf{q}},\tilde{\mathbf{s}})$ values. The less "fit" vectors $\tilde{\mathbf{q}}$ are replaced by new randomly generated individuals. This procedure is repeated iteratively until χ^2 reaches a predefined tolerance or the number of iterations (generations) reaches its predefined maximum. Finally, one obtains the coordinates of a probable minimum value of $\chi^2(\mathbf{q},\mathbf{s})$ and its coordinates $\mathbf{q}_m$ and $\mathbf{s}_m$ in the space of the allowed values of $\mathbf{q}$ and $\mathbf{s}$. In the above optimization scheme, the GA and MLR routines converge simultaneously and both sets of the model parameters are optimized in parallel. By running the GA-MLR optimizer several times (each time with a different seed value for the random number generator), one recovers statistical information on the probable coordinates (a distribution of the estimates $\mathbf{q}_m$ and $\mathbf{s}_m$) of the global minimum of χ^2. If it is required, one can use the values of the fitted parameters returned by the GA optimizer as the initial guesses for the GE optimization method, as the final step in the optimization procedure.

It is very essential to stress here that the solution (8) to (6) has rather a formal character and a highly recommended method for solving of the normal equations is the Gauss−Jordan elimination method.[19] To avoid the possible difficulties arising from the roundoff errors (when solving the normal equations) or from the normal equations being very close to singular, the singular value decomposition (SVD) is the most recommended method for minimization of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ with respect to vector $\mathbf{s}$ of the combination coefficients a_k in eq 3 (see ref 19 for very detailed discussion).

3. Illustrative Example. Fluorescence Decay Surface in the Two-Excited-State Interconversion Problem

To exemplify the GA-MLR optimization method, we here consider one of the simplest cases, i.e., the biexponential fluorescence decay surface emitted by a compound undergoing two-excited-state interconversion process (e.g., intramolecular charge transfer (ICT), twisted intramolecular charge transfer (TICT) or intramolecular proton transfer (PT) process), controlled by four kinetic rate constants k_1, k_{12}, k_2, and k_{21}, as is shown in Scheme 1. We here want to demonstrate the simplicity and very high efficiency of the GA-MLR optimization method. Additionally, we want to show that the GA-MLR optimizer, when used together with the efficient methods for the analysis of the kinetic fluorescence decays for the ICT, TICT, and PT processes, outlined in ref 23, allows for almost immediate recovery of the emission bands and the values of the rates of state-to-state kinetic relaxation assumed in the simulations.

To make the discussion a bit more general, we begin with the application of the GA-MLR optimization approach to the case of a single multiexponential fluorescence decay. Let us suppose that $I_m(t)$ represents the histogram of an experimentally recovered multiexponential fluorescence decay and that $I_c(\mathbf{q},\mathbf{s},t)$ is the model fluorescence decay at δ-pulse excitation

$$I(\mathbf{q},\mathbf{s},t) = \sum_{k=1}^{M} \alpha_k^{(\delta)} \exp(-tL_k) = \sum_{k=1}^{M} \alpha_k^{(\delta)} f_k^{(\delta)}(t) \qquad (9)$$

paramatrized by the vector of decay rates $\mathbf{q} = [L_1, ..., L_M]$ (nonlinear parameters) and the vector of corresponding amplitudes $\mathbf{s} = [\alpha_1^{(\delta)}, ..., \alpha_M^{(\delta)}]$ (linear parameters). In this case eq 2 becomes[20]

$$\chi^2(\mathbf{q},\mathbf{s}) = \frac{1}{\nu_f} \sum_{i=1}^{N} \frac{(I_m(i) - I_c(\mathbf{q},\mathbf{s},i))^2}{I_m(i)} \qquad (10)$$

where $I_c(\mathbf{q},\mathbf{s},i)$ is the histogram of $I(\mathbf{q},\mathbf{s},i)$ convoluted with the δ-pulse-excitation instrument response function $I_r(i)$ and $\sigma_i = \sqrt{I_m(i)}$ are the experiment errors within the Poisson statistics.

The GA optimizer generates randomly a trial population of the complete sets of the decay rates (i.e., a trial set of vectors $\tilde{\mathbf{q}}$), which take the values from the intervals bounded by their predefined lower and upper limits, and which are used in the evaluation of the corresponding convolution integrals

$$f_k(t) = \int_{-\infty}^{t} f_k^{(\delta)}(t-t')\, I_r(t')\, dt' \qquad (k = 1, ..., M) \quad (11)$$

For each trial vector $\tilde{\mathbf{q}}$ of the decay rates one has to minimize the corresponding $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ function of $\mathbf{s}$ at fixed $\tilde{\mathbf{q}}$, i.e.

$$\chi^2(\tilde{\mathbf{q}},\mathbf{s}) = \frac{1}{\nu_f} \sum_{i=1}^{N} \frac{1}{I_m(i)} \left(I_m(i) - \sum_{k=1}^{M} \alpha_k f_k(i)\right)^2 \qquad (12)$$

The minimum-value condition of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ with respect to $\mathbf{s}$, which is the vector of the amplitudes α_k standing at corresponding convoluted histograms $f_k(t)$ in $I_c(\mathbf{q},\mathbf{s},t)$, leads, according to eq 4, to the following set of normal equations in the MLR optimization

$$\frac{\partial \chi^2(\tilde{\mathbf{q}},\mathbf{s})}{\partial \alpha_j} = \sum_{k=1}^{M} \left(\sum_{i=1}^{N} \frac{f_j(i)\, f_k(i)}{I_m(i)} \right) \alpha_k - \sum_{i=1}^{N} f_j(i) = 0 \qquad (13)$$

where $j = 1, ..., M$, which is equivalent to the matrix equation (6), and where

$$A_{jk} = \sum_{i=1}^{N} \frac{f_k(i)\, f_j(i)}{I_m(i)} \qquad c_j = \sum_{i=1}^{N} f_j(i) \qquad (14)$$

A formal solution to eq 13 is given by the formula (8). Substituting the recovered values of the amplitudes α_k into eq 12 and repeating this procedure over all trial vectors $\tilde{\mathbf{q}}$, the trial set of the values of $\chi^2(\tilde{\mathbf{q}},\tilde{\mathbf{s}})$ is obtained, which, together with the trial vector $\tilde{\mathbf{q}}$, is subjected to further steps within the GA algorithm, as was discussed in the previous section.

In the case of a globally analyzed set of multiexponential decays, the GA optimizer modifies the nonlinear parameters simultaneously in all decays, according to the assumed linking scheme, and the linear parameters are evaluated by means of the MLR method for each individual decay.

For the two-excited-state problem depicted in Scheme 1, the steady-state emission bands PS and TS and time-resolved fluorescence decay surface were simulated for the following values of the kinetic rate constants: $k_1^{-1} = 5$ ns, $k_{12}^{-1} = 0.5$ ns, $k_2^{-1} = 4.7$ ns and $k_{21}^{-1} = 1$ ns. The normalized shapes of the spectral distributions of the photons emitted from both states,

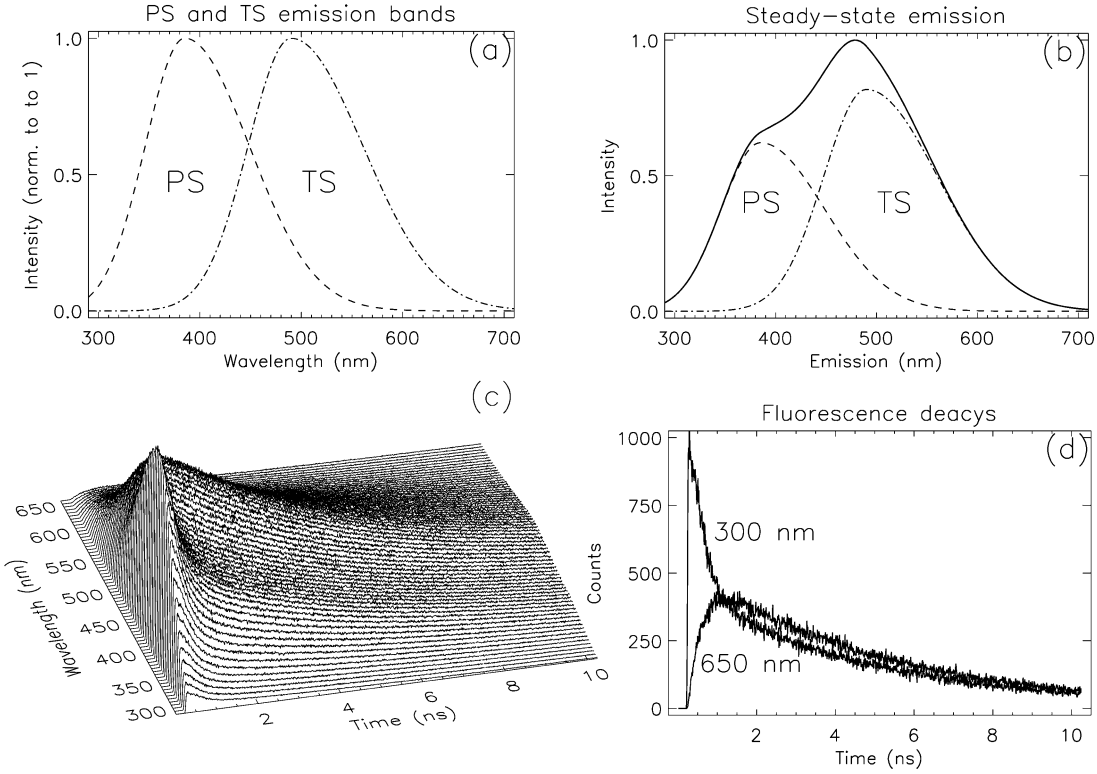


Figure 1. Simulated data: (a) assumed spectral shapes of PS and TS emission bands; (b) simulated steady-state PS, TS and effective emission bands; (c) simulated fluorescence decay surface; (d) two examples of simulated fluorescence decays, for $\lambda_{em} = 300$ nm and $\lambda_{em} = 650$ nm.

used in the simulations, are shown in Figure 1a. With these data we calculated the decay rates $L_1^{-1} = 312$ ps and $L_2^{-1} = 4796$ ps, and the amplitudes $\alpha_1(\lambda_{em})$ and $\alpha_2(\lambda_{em})$ for the emission wavelengths from 300 to 650 nm, in the step of 5 nm (see ref 23). The calculated intensities of steady-state bands PS and TS,[23] normalized to unity with respect to the maximum value of the effective fluorescence emission, are shown in Figure 1b. The obtained decay rates L_1 and L_2 and 71 pairs of the amplitudes $\alpha_1(\lambda_{em})$ and $\alpha_2(\lambda_{em})$ were used in generating the synthetic fluorescence decay surface, which is composed of 71 biexponential fluorescence decays (9) convoluted with the experimentally recovered histogram of the scattered laser pulse collected in 1024 channels (channel width 10 ps). The synthetic fluorescence decay surface was normalized to set the counts number at its maximum to 2×10^4 counts. Different Poisson noises were added to different constituent fluorescence decays. The obtained fluorescence decay surface and the constituent fluorescence decays, for $\lambda_{em} = 300$ nm and $\lambda_{em} = 650$ nm, are shown in Figure 1c,d. The synthetic fluorescence decay surface involves 144 decay parameters, from which 2 decay rates L_1 and L_2 are fitted by the GA optimizer and 142 amplitudes are calculated from the MLR method.

In the analysis of the synthetic data the population of trial vectors $\mathbf{q} = [L_1, L_2]$ was set to 100 individuals and the predefined number of generations was set to 50. The GA-MLR optimizer was run 50 times, each time with a different seed value for the random number generator. The following upper and lower bounds for the fitted parameters were set: $L_1^{-1} = [100, 1000]$ ps and $L_2^{-1} = [1100, 8000]$ ps. The GA-MLR optimizer implements the PIKAIA subroutine.[24]

The final results of 10 different runs of the GA-MLR optimizer are presented in Figure 2a−d, where the histograms of χ^2 converge to the values very close to unity (Figure 2a) and the histograms of the decay rates (L_1 and L_2) converge to the values of the decay times almost identical to those obtained in the simulation of the data (Figure 2b,c). The same deals also with the recovered amplitudes, shown in Figure 2d. The estimates of the decay rates L_1 and L_2 recovered in all 50 runs of the GA-MLR optimizer are displayed in Figure 2e,f. They are very closely distributed around the simulated values and their averaged values, calculated from the series of all 50 values returned by GA-MLR, are $L_1^{-1} = 311$ (±4) ps and $L_2^{-1} = 4796$ (±6) ps. The values of $L_1^{-1} = 312$ (±5) ps and $L_2^{-1} = 4798$ (±13) ps we obtained from the GE analysis of the data, in which the averaged values of all model parameters estimated from GA-MLR optimizer were assumed as the initial guesses in the GE optimization.

We here want to demonstrate that the combination of the GA-MLR optimizer with the algorithm outlined in ref 23 enables us to perform a fast separation of the constituent stead-state emission bands and decay surfaces of the excited-states PS and TS and enables us to estimate the values of the kinetic rate constants. The algorithm outlined in ref 23 is not a general method and applies to particular experimental cases in which the emission from the PS state can be detected separately and can be easily verified experimentally[23] (this is also displayed in the example). In a more general case, in which the PS and TS emission bands totally overlap, the more general numerical treatments have to be applied (see the references cited in ref 23; see also ref 26 and the references therein). According to the algorithm outlined in ref 23 (see also ref 25), from the sums and differences for the decay parameters, i.e., $L_1 + L_2$, $L_1 - L_2$, $\alpha_1(\lambda_{em}) + \alpha_2(\lambda_{em})$ and $\alpha_1(\lambda_{em}) - \alpha_2(\lambda_{em})$, the following relations can be derived:

$$U_1 = \frac{\alpha_1(\lambda_{em})L_1 + \alpha_2(\lambda_{em})L_2}{\alpha_1(\lambda_{em}) + \alpha_2(\lambda_{em})} + \frac{X(\lambda_{em})}{\alpha_1(\lambda_{em}) + \alpha_2(\lambda_{em})} \quad (15)$$

$$U_2 = \frac{\alpha_1(\lambda_{em})L_2 + \alpha_2(\lambda_{em})L_1}{\alpha_1(\lambda_{em}) + \alpha_2(\lambda_{em})} - \frac{X(\lambda_{em})}{\alpha_1(\lambda_{em}) + \alpha_2(\lambda_{em})} \quad (16)$$

$$W = \frac{1}{4}[(L_1 - L_2)^2 - (U_1 - U_2)^2] \qquad (17)$$

where the "true" (emission-wavelength-independent) values of parameters U_1, U_2 and W are defined by the appropriate combinations of the four kinetic rates, i.e.

$$U_1 = k_1 + k_{12} \qquad U_2 = k_2 + k_{21} \qquad W = k_{12}k_{21} \quad (18)$$

Note that U_1^{-1} and U_2^{-1} are the fluorescence lifetimes of both excited states.

At the emission wavelengths λ_{PS}, at which the emission from state TS does not contribute to the detected fluorescence signal, $X(\lambda_{PS}) = 0$ and the right-hand sides of eqs 5−17 take constant values, which are the estimates of "true" values of U_1, U_2 and W, defined by (18). According to Figure 3a,b, in the range of the emission wavelengths 300−360 nm, the fluorescence signal from the TS state does not contribute or this contribution is negligible, and we obtain the following estimated "true" values of U_1, U_2 and W: $U_1^{-1} \simeq 453$ (±4) ps, $U_2^{-1} \simeq 825$ (±7) ps and $W \simeq 2.01$ (±0.04) $\times 10^{-6}$ ps^{-2} . They are in agreement with the corresponding values obtained in the simulations, i.e., $U_1^{-1} \simeq 455$ ps, $U_2^{-1} \simeq 825$ ps and $W \simeq 2.00 \times 10^{-6}$ ps^{-2}.

As was discussed in ref 23 (see also ref 25), in some practical cases, the application of the approximations resulting from the "high-temperature" (HT) and "low-temperature" (LT) kinetic limits, enables us to reduce the number of the kinetic rates in eqs 18. The parameters identifying the HT and LT limits

$$R_{HT} = \frac{U_1\,\alpha_2(\lambda_{PS})}{U_2\,\alpha_1(\lambda_{PS})} \qquad R_{LT} = \frac{U_1\,L_2}{U_2\,L_1} \qquad (19)$$

in our case take the values $R_{HT} = 0.91$ and $R_{LT} = 0.11$, and the closure property $R_{HT} + R_{LT} \simeq 1$ holds. This result means that, although the values of the kinetic rates assumed in the simulations do not correspond perfectly to the HT kinetic limit (in such cases $R_{HT} \simeq 1$ and $R_{LT} \simeq 0$), the first relation in the HT kinetic limit is fulfilled to the first approximation, i.e., $k_1 \ll k_{12}$.[27] Hence, only three kinetic rates take the nonvanishing values in eqs 18, and consequently, one can assume the values $k_{12}^{-1} \simeq U_1^{-1} = 453$ ps, $k_{21}^{-1} \simeq (W/k_{12})^{-1} = 1.1$ ns and $k_2^{-1} \simeq (U_2 - k_{21})^{-1} = 3.3$ ns, according to (18), which are in agreement (to the first approximation) with the values of the same parameters assumed in the simulations. Very many experimental examples, better displaying how very useful can be the application of temperature-dependent kinetic LT and HT approximations in the analysis of ICT, TICT and PT processes, are demonstrated and discussed in a very systematic way in refs 23 and 25.

Having estimated the values of U_1 and U_2, the function $X(\lambda_{em})$ can be reconstructed from relations 15 and 16, i.e.

$$X(\lambda_{em}) = \frac{1}{2}[(U_1 - U_2)(\alpha_1(\lambda_{em}) + \alpha_2(\lambda_{em})) - (L_1 - L_2)(\alpha_1(\lambda_{em}) - \alpha_2(\lambda_{em}))] \quad (20)$$

where U_1 and U_2 take the estimated values indicated in Figure 3a. The plot of $X(\lambda_{em})/(L_1 - L_2)$ reflects the spectral shape of the emission band TS[23] (seen in Figure 3c). Knowing $X(\lambda_{em})$, one can decompose the fluorescence decays surface $I(\lambda_{em},t)$ (see Figure 1c) and effective steady-state emission band $I^{ss}(\lambda_{em})$ (see Figure 1b) into two constituent, PS and TS, decay surfaces and steady-state emission bands.[23] For example, the steady-state intensities of the emission bands TS and PS, i.e., $I_{PS}^{ss}(\lambda_{em})$ and $I_{TS}^{ss}(\lambda_{em})$, can be obtained from

$$I_{TS}^{ss}(\lambda_{em}) = C_{TS}(\lambda_{em})\, I^{ss}(\lambda_{em})$$
$$I_{PS}^{ss}(\lambda_{em}) = I^{ss}(\lambda_{em}) - I_{TS}^{ss}(\lambda_{em}) \qquad (21)$$

where $C_{TS}(\lambda_{em})$ expresses the contribution of $I_{TS}^{ss}(\lambda_{em})$ to the measured effective intensity $I^{ss}(\lambda_{em})$, and[23]

$$C_{TS}(\lambda_{em}) = \frac{X(\lambda_{em})}{X(\lambda_{em}) + (\alpha_1(\lambda_{em}) + \alpha_1(\lambda_{em}))\,U_2} \qquad (22)$$

The coefficients $C_{TS}(\lambda_{em})$ and the constituent steady-state emission bands, calculated for the recovered decay parameters, are shown in Figure 3d,e, correspondingly. The recovered emission bands $I_{PS}^{ss}(\lambda_{em})$ and $I_{TS}^{ss}(\lambda_{em})$ perfectly correspond to the simulated bands shown in Figure 1b.

4. Optimization of Linear, Weakly Nonlinear and Nonlinear Model Parameters by Means of the GA-NR Optimizer

In this section we come back to the GA-NR algorithm introduced in our recent publication[13] and mentioned in short in the Introduction. We here want to add a few important explanatory comments on the application of the GA-NR optimizer to linear and weakly nonlinear model parameters that occur in the same optimization problem together with the nonlinear parameters.

Very recently, the authors of the ref 28−30 have presented an excellent application of the GA optimization method to the automated assignment of high-resolution spectra. By introducing several improvements, they have obtained a GA-based optimization method enabling very fast analysis of dense spectra in which very many vibronic spectra of conformers and isoto-

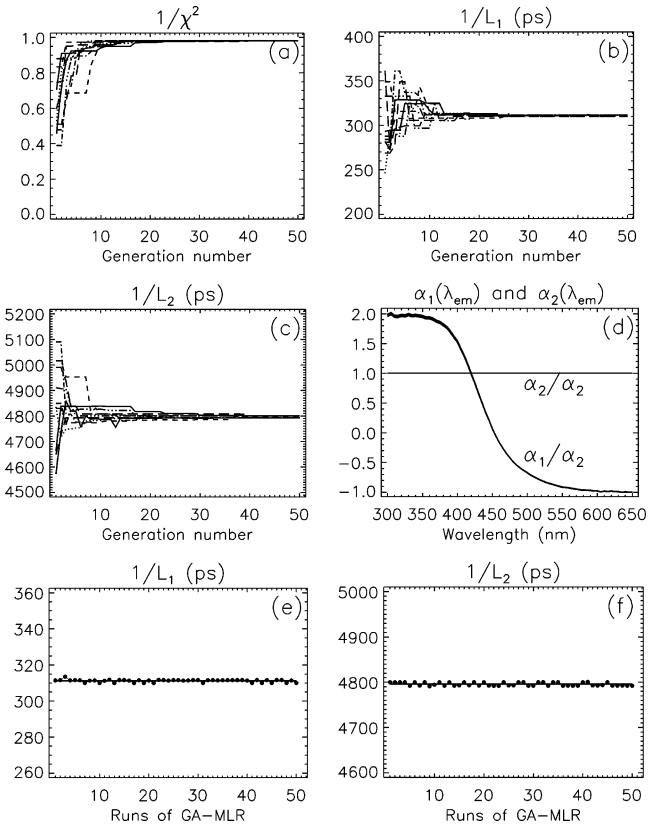


Figure 2. Results of the GA-MLR optimization: (a)−(c) 10 examples of the registrated histograms of $\chi^2(\tilde{\mathbf{q}}, \tilde{\mathbf{s}})$ and both decay rates L_1 and L_2; (d) 10 examples of recovered amplitudes $\alpha_1(\lambda_{em})$ and $\alpha_2(\lambda_{em})$; (e)−(f) recovered values of L_1 and L_2 in all 50 runs of the GA-MLR optimizer.

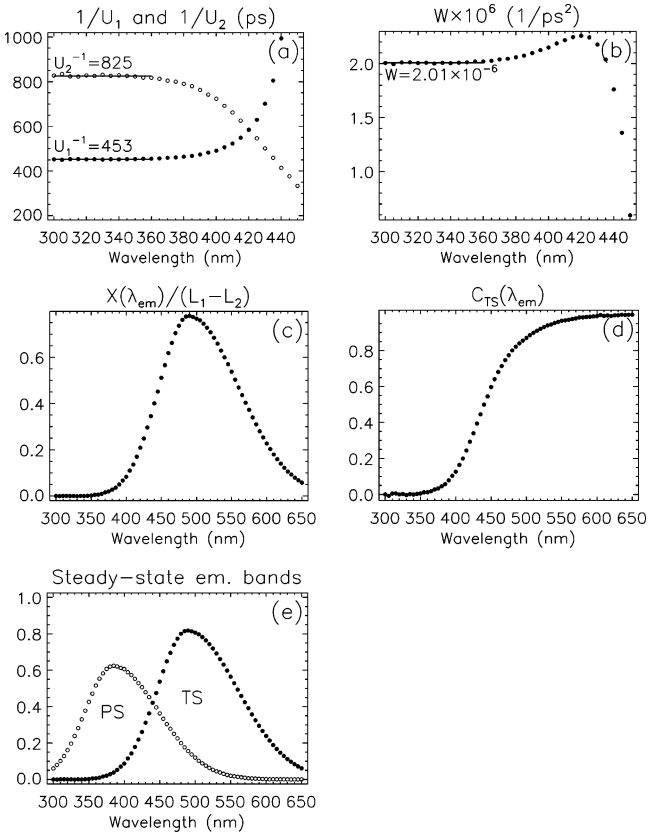


Figure 3. Analysis of the decay parameters: (a, b) estimated values of U_1, U_2 and W; (c) reconstructed spectral shape of TS emission band; (d) coefficient $C_{TS}(\lambda_{em})$; (e) decomposed steady-state emission bands PS and TS.

pomrers overlap, and for strongly congested spectra in dimer systems. To accelerate the assigned fit calculations, i.e., to avoid using of the time-consuming Levenberg−Marquardt algorithm, in ref 30 the authors have outlined the optimization method in which the minimum-value condition for χ^2 is expanded in the Taylor series and the model parameters are recovered iteratively. This method is not coupled with the GA optimizer. From the mathematical point of view, the method described in ref 30 and the NR method are equivalent, and this is easily seen after skiping all simplifying assumptions made in the Appendix B of ref 30.

If vector **s** represents the set of linear parameters and if $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ is quadratic function of **s** (i.e., the problem is linear; model function is linear combination of nonlinear functions), the minimum-value condition (4) for $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ leads to the set of normal equation and the best estimates of linear parameters (i.e., vector $\tilde{\mathbf{s}}$) are obtained from the MLR method, as was discussed in section 2.

If $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ is not quadratic function of **s** (i.e., the problem is nonlinear), vector $\tilde{\mathbf{s}}$ can be recovered from the minimum-value condition (4) for the quadratic approximation of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ and from the Newton−Raphson algorithm.[17] Vector **s** may represent solely linear parameters (i.e., the model function is a multi-linear combination of nonlinear functions), solely weakly nonlinear parameters or a set of partly linear and partly weakly nonlinear parameters.

The quadratic approximation of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ with respect to $(\mathbf{s}-\mathbf{s}^{(j)})$ is given by eq 5 in ref 13 and has been obtained from the Taylor series expansion of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ about $\mathbf{s}^{(j)}$ at fixed vector $\tilde{\mathbf{q}}$ of nonlinear parameters returned temporarily by the GA optimizer. The minimum-value condition (4) for the quadratic approximation of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ leads to

$$-\left.\frac{\partial^2\chi^2(\tilde{\mathbf{q}},\mathbf{s})}{\partial\,\mathbf{s}^2}\right|_{\mathbf{s}^{(j)}}(\mathbf{s}-\mathbf{s}^{(j)})=\left.\frac{\partial\chi^2(\tilde{\mathbf{q}},\mathbf{s})}{\partial\,\mathbf{s}}\right|_{\mathbf{s}^{(j)}} \tag{23}$$

In the Newton−Raphson algorithm, if $\mathbf{s}_0$ is an initial guess of $\tilde{\mathbf{s}}$, the sequential calls of the following recursion formula[13,17]

$$\mathbf{s}^{(j+1)}=\mathbf{s}^{(j)}+\left(-\left.\frac{\partial^2\chi^2(\tilde{\mathbf{q}},\mathbf{s})}{\partial\mathbf{s}^2}\right|_{\mathbf{s}^{(j)}}\right)^{-1}\left[\left.\frac{\partial\chi^2(\tilde{\mathbf{q}},\mathbf{s})}{\partial\mathbf{s}}\right|_{\mathbf{s}^{(j)}}\right] \tag{24}$$

will results in a sequence of $\mathbf{s}^{(1)}$, $\mathbf{s}^{(2)}$, ... that converges to vector $\tilde{\mathbf{s}}$.

The recursion formula (24), for the explicit form of $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ given by eq 2, takes the form

$$\mathbf{s}^{(j+1)}=\mathbf{s}^{(j)}+\mathbf{A}^{-1}\mathbf{c} \tag{25}$$

where

$$A_{mn}=\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\left[\left.\frac{\partial y(\tilde{\mathbf{q}},\mathbf{s},x_i)}{\partial s_m}\right|_{\mathbf{s}^{(j)}}\left.\frac{\partial y(\tilde{\mathbf{q}},\mathbf{s},x_i)}{\partial s_n}\right|_{\mathbf{s}^{(j)}}-(y_i-y(\tilde{\mathbf{q}},\mathbf{s}^{(j)},x_i))\left.\frac{\partial^2 y(\tilde{\mathbf{q}},\mathbf{s},x_i)}{\partial s_m\partial s_n}\right|_{\mathbf{s}^{(j)}}\right] \tag{26}$$

and

$$c_m=\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(y_i-y(\tilde{\mathbf{q}},\mathbf{s}^{(j)},x_i))\left.\frac{\partial y(\tilde{\mathbf{q}},\mathbf{s},x_i)}{\partial s_m}\right|_{\mathbf{s}^{(j)}} \tag{27}$$

The term $\Delta\mathbf{s}=\mathbf{A}^{-1}\mathbf{c}$, which represents the vector of corrections modifying vector $\mathbf{s}^{(j)}$ into new vector $\mathbf{s}^{(j+1)}$, has rather a formal form, and similar to the case of the MLR optimizer, the most recommended methods for obtaining $\Delta\mathbf{s}$ are the Gauss−Jordan elimination or SVD method applied to the system of coupled linear equations (23), at each iteration in the NR algorithm.

The conditions (4) and (23) define, in principle, the multiple minima, maxima and saddle points of χ^2 function, in a general case. For linear problems they define just the minimum of χ^2. For nonlinear problems, if NR method is used independently (though not as the inner optimizer in GA-NR), one has to make sure that the results obtained from both conditions really represent the global minimum of χ^2, though not one of the local minima, maxima or one of the saddle points. In Figure 4 we show schematically a χ^2 function of the model parameters α, which contains multiple minima (min), maxima (max) and saddle points (s), and for which the conditions (4) and (23) hold.

From the mathematical point of view, these three categories of the recovered values of model parameters can be distinguished from the values of the second-order derivatives of χ^2. From the point of view of reliable numerical recovery of the most probable global minimum of χ^2, the data should be analyzed according to the numerical strategy suggested and tested in ref 12 for the case of the global analyses of polarized fluorescence decays for the case of membrane vesicles and planar (macroscopically ordered) membranes, by means of the GE optimizer, namely, (a) the analysis is repeated for very many sets of initial guesses for the fitted parameters (they are scanned between the reasonably defined upper and lower limits) and (b) the obtained set of χ^2 values is sorted out in sequence between their maximum and minimum values. Such strategy enables us to obtain a revealing insight into the multidimensional distribution of the χ^2 values fulfilling the conditions (4) and (23), projected onto the axes corresponding to the adjustable

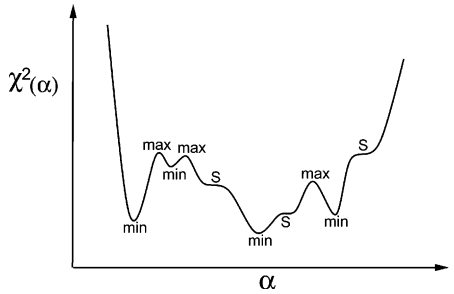


Figure 4. Schematic presentation of $\chi^2(\alpha)$ function possessing multiple minima, maxima and the saddle points.

parameters. Consequently, the most probable coordinates of the global minimum of χ^2 can be obtained. In the case of NR optimization method, the recovered set of χ^2 contains the values of χ^2 that relate to multiple minima, maxima and saddle points. This is in contrast to GE optimizer because this optimizer ignores all maxima and saddle points of the χ^2 surface. Therefore, a complete and realiable data analysis by means of the NR algorithm requires much denser grid of the initial guesses for the fitted parameters, as compared to the GE optimizer. Below we discuss this point for the case of the GA-NR optimizer, in which NR algorithm is the inner optimizer controlled by the outer one GA.

If the NR method is embedded in the GA optimizer (i.e., NR is the inner optimizer in the GA-NR method), the GA optimizer provides a trial population of vectors $\tilde{\mathbf{q}}$ (i.e., a trial population of the complete sets of nonlinear parameters in $y(\tilde{\mathbf{q}},\mathbf{s},x)$) and also a trial population of the complete sets of initial guesses $\mathbf{s}_0$ of linear and/or weakly nonlinear parameters $\mathbf{s}$. Each trial vector $\tilde{\mathbf{q}}$ returns the corresponding $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ function of vector $\mathbf{s}$. For the predefined number of iterations to be performed by the NR algorithm, the correspondig set of vectors $\tilde{\mathbf{s}}$ is obtained from the optimization based on the recursion formula (25). Once both populations of vectors $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{s}}$ are obtained, the coresponding set of $\chi^2(\tilde{\mathbf{q}},\tilde{\mathbf{s}})$ values is known automatically and is subjected to further steps of the GA optimizer, as was outlined in section 2.

The application of NR optimizer to linear or weakly nonlinear parameters may accelerate the convergence of the GA method, as distinct from the case when all model parameters are treated as if they were nonlinear. It is expected that in such cases the GA-NR optimizer will converge at much smaller populations of the complete sets of the model parameters and a much smaller number of generations will be performed by GA optimizer. The application of this method to the model parameters of an effectively higher "degree" of the nonlinearity, or just to nonlinear parameters, will lead to much slower convergence of the GA-NR optimizer as compared to the case when they are optimized directly by the outer optimizer (GA).

In the above proposed algorithm, the initial guesses $\mathbf{s}_0$ are selected by the GA optimizer (in principle they are fitted by GA) because, in a general case, an external prediction of their right values may represent a serious difficulty. Furthermore, selection of the initial guesses for $\mathbf{s}$ by the GA algorithm minimizes the risk of trapping the convergence of the optimization process at a local minimum of χ^2 with respect to $\mathbf{s}$ because, in this case, a greater number of initial guesses $\mathbf{s}_0$ is taken into account simultaneously in the GA optimization method. For the same reason, the recovered "best" estimates $\mathbf{s}_m$ will describe just the minimum of χ^2 with respect to $\mathbf{s}$, though not one of the local maxima or one of the saddle points. In other words, the GA optimizer, by testing and modifying simultaneously a greater number of vectors $\mathbf{s}_0$ (with the same parallel process for vectors $\mathbf{q}$), drives the optimization toward the global minimum of χ^2 with respect to both sets of the fitted parameters.

Very interesting is the case when $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ is the nonquadratic function of linear parameters $\mathbf{s}$. As was mentioned at the begining of this section, such cases occur for the model functions which are multilinear combinations of nonlinear functions. Let us consider one of the simplest examples of such a class of the model functions, namely,

$$y(\tilde{\mathbf{q}},\mathbf{s},x) = abf_1(\tilde{\mathbf{q}}_1,x) + acf_2(\tilde{\mathbf{q}}_2,x) + bcf_3(\tilde{\mathbf{q}}_3,x) \quad (28)$$

where $\boldsymbol{s} = [a, b, c]$, and $\tilde{\mathbf{q}}_1$, $\tilde{\mathbf{q}}_2$ and $\tilde{\mathbf{q}}_3$ are the vectors of nonlinear parameters returned temporarily by the GA optimizer (they contain the model parameters that occur in vector $\tilde{\mathbf{q}}$). $y(\tilde{\mathbf{q}},\mathbf{s},x)$ depends linearly on each parameter in $\mathbf{s}$, individually, but $\chi^2(\tilde{\mathbf{q}},\mathbf{s})$ is not a quadratic function of $\mathbf{s}$. It is important to stress here that for weakly linear parameters $\mathbf{s}$ all the second-order derivatives in eq 26 take the nonzero values, in a general case. However, if $\mathbf{s}$ represents the set of linear parameters, $\partial^2 y(\tilde{\mathbf{q}},\mathbf{s},x_i)/\partial s_m \partial s_n = 0$ when $m = n$, and $\partial^2 y(\tilde{\mathbf{q}},\mathbf{s},x_i)/\partial s_m \partial s_n \neq 0$ if $m \neq n$, in a general case. As will be seen below, for the model functions that are linear combinations of nonlinear functions, all second-order partial derivatives take the zero values in eq 26. This means that the second term in eq 26 distinguishes between the linear, multilinear and (strongly or weakly) nonlinear model functions of $\mathbf{s}$.

If vector $\mathbf{s}$ represents the set of linear parameters on which χ^2 depends quadratically, i.e., the model function is given by linear combination of nonlinear functions (see eq 1), the second-order partial derivative in (26) disappears, i.e., $\partial^2 y(\tilde{\mathbf{q}},\mathbf{s},x_i)/\partial s_m \partial s_n = 0$, and expression for the matrix elements A_{mn} simplifies to

$$A_{mn} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left.\frac{\partial y(\tilde{\mathbf{q}},\mathbf{s},x_i)}{\partial s_m}\right|_{\mathbf{s}(j)} \left.\frac{\partial y(\tilde{\mathbf{q}},\mathbf{s},x_i)}{\partial s_n}\right|_{\mathbf{s}(j)} \quad (29)$$

and the expression for c_m remains unchanged.

Let us assume that $y(\tilde{\mathbf{q}},\mathbf{s},x)$ contains a single linear parameter a, i.e.

$$y(\tilde{\mathbf{q}},\mathbf{s},x) = aF(\tilde{\mathbf{q}},x) \quad (30)$$

With this assumption, $\mathbf{A}$ and $\mathbf{c}$ become the scalars A and c, and

$$A^{-1}c = \frac{\sum_{i=1}^{N} 1/\sigma_i^2\, y_i\, F(\tilde{\mathbf{q}},x_i)}{\sum_{i=1}^{N} 1/\sigma_i^2 F(\tilde{\mathbf{q}},x_i)^2} - a^{(j)} \quad (31)$$

Consequently, the recursion formula (25) becomes

$$a^{(j+1)} = a^{(j)} + A^{-1}c = \frac{\sum_{i=1}^{N} 1/\sigma_i^2\, y_i\, F(\tilde{\mathbf{q}},x_i)}{\sum_{i=1}^{N} 1/\sigma_i^2 F(\tilde{\mathbf{q}},x_i)^2} \quad (32)$$

From the above obtained recursion formula two very important conclusions can be drawn. First, for linear parameters the NR algorithm converges after the first iteration (i.e., any further iteration returns the same value of the recovered linear parameter because the value of $a^{(j+1)}$ does not depend on the former value $a^{(j)}$, according to (32)). Second, the values of initial guesses for the linear parameters may be chosen in an arbitrary way because

they do not effect the values of the optimized linear parameters in the NR algorithm (i.e., $a^{(j+1)}$ does not depend on $a^{(j)}$ begining with the first iteration). These conclusions have general meaning and they hold for a general form of $y(\tilde{\mathbf{q}},\mathbf{s},x)$ given by eq 1. As is seen from the above, for the optimization problems in which χ^2 depends quadratically on linear paremeters **s**, the GA-NR and GA-MLR optimizers are equivalent.

5. Discussion

In this article we have discussed and exemplified the combined genetic algorithms (GA) and multiple linear regression (MLR) optimization approach (GA-MLR) that applies to nonlinear optimization problems, in which the nonlinear model functions are linear combinations of nonlinear functions. As was demonstrated, the GA-MLR optimizer exploits all advantages of the genetic algorithms optimization method and is based on a simple numerical trick consisting of appropriate combining of two well-known optimization methods. GA-MLR optimizes the nonlinear and linear model parameters in parallel, i.e., the MLR optimizer is embedded in the GA one. The GA-MLR approach involves the idea of separability of nonlinear and linear fitted parameters, introduced, for the first time, by the variable projection (VP) formalism,[21] and which has a very advanced and very strong mathematical background. The GA-MLR (likewise GA-FOD and GA-NR introduced in ref 13) algorithm is an intuitive treatment in which the multiple linear least-squares method is the only one strictly mathematical "tool" employed. VP and GA-MLR are two very different treatments to the same class of nonlinear least-squares problems. The VP optimization method has already been verified on very many optimization problems.[22] The GA-MLR method is a newly introduced algorithm and, apart from the numerical tests demonstrated in this article (and in ref 13 for a more complicated problem in which the nonlinear function is obtained from the numerical solution to the potential-restricted diffusion equation), it has not been tested yet on real experimental data. However, the analysis of a synthetic fluorescence decay surface for the two-excited-state interconversion process, depicted in Scheme 1, has displayed the simplicity and serious advantages of the GA-MLR optimizer. This example provides some indications that enable us to expect that the GA-MLR approach may find interesting applications in some practical optimization problems. It seems, for example, that the combination of the GA-MLR optimization method with the algorithm outlined in ref 23, and designed for the analysis of fluorescence decays of the compounds undergoing the ICT, TICT and PT processes, may be a very promising tool in some practical cases of photochemical studies of the compounds undergoing the excited-state process of relatively simple character, to which the discussed treatment may apply.

We have added a few important explanatory comments on the application of the GA-NR to the optimization problems in which the linear and/or weakly nonlinear model parameters, fitted by the NR method, occur together with the nonlinear ones being recovered by the GA optimizer. As was indicated, the GA-NR method applies also to nonlinear problems in which the model functions are multilinear combinations of nonlinear functions. The properties of the GA-NR optimizer will be displayed, in a more systematic and more detailed way, elsewhere.

It has to be noted here that the GA-MLR and GA-NR optimization methods are not the universal algorithms. Genetic algorithms, likewise the gradient expansion optimization method, apart from their intersting advantages, possess also some disadvantages. For example, for particular shapes of the χ^2 surface in the space of fitted model parameters, the GA optimizer converges very slowly. This happens in most of the cases when this surface is very flat, and thereby, the convergence of GA is very hardly achieved. In all such cases the GE optimization method is much more efficient. In such cases GA-MLR and GA-NR optimizers can be used for obtaining a set of "good" initial guesses for the GE optimization.

Acknowledgment. This work was supported by the Rector of the Nicolaus Copernicus University, under grant 504-F. This paper is dedicated to the memory of my dear daughter Karolina.

References and Notes

(1) Knutson, J. R.; Beechem, J. M.; Brand, L. *Chem. Phys. Lett.* **1983**, *102*, 501−507.
(2) Beechem, J. M.; Ameloot M.; Brand, L. *Anal. Instrum.* **1985**, *14*, 379−402.
(3) Beechem, J. M. *Chem. Phys. Lipids* **1989**, *50*, 237−251.
(4) Beechem, J. M. *Methods Enzymol.* **1992**, *210*, 37−54.
(5) Holzwarth, A. R. *Methods Enzymol.* **1995**, *246*, 334−362.
(6) Holzwarth, A. R. Data analysis of time-resolved measurements. In *Biophysical Techniques in Photosynthesis*; Amesz, J., Hoff, A. J., Eds.; Kluwer Academic Publishing: Dordrecht, The Netherlands, 1996; pp 75−92.
(7) Dioumaev, A. K. *Biophys. Chem.* **1997**, *67*, 1−25.
(8) Löfroth, J. E. *J. Phys. Chem.* **1986**, *90*, 1160−1168.
(9) Brochon, J. C. *Methods Enzymol.* **1994**, *240*, 262−311.
(10) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Kluwer Academic/Plenum Publishers: New York, Boston, Dordrecht, London, Moscow, 1999.
(11) Valeur, B. *Molecular Fluorescence*; Wiley-VCH: Weinheim, 2002.
(12) Fisz, J. J. *Chem. Phys. Lett.* **2005**, *407*, 1−7.
(13) Fisz, J. J.; Buczkowski, M.; Budziński, M. P.; Kolenderski, P. *Chem. Phys. Lett.* **2005**, *407*, 8−12.
(14) Holland, J. H. *J. Assoc. Comput. Math.* **1962**, *3*, 297−319.
(15) Goldberg, D. A. *Genetic Algorithms in Search of Optimization and Machine Learning*; Addison-Wesley: Reading, MA, 1989.
(16) Davis, L. *Handbook of Genetic Algorithms*; Van Nostrand Reinhold: New York, 1991.
(17) Tanner, M. A. *Tools for Statistical Inference. Methods for the Exploration of Posterior Distributions and Likelihood Functions*; Springer-Verlag: New York, Berlin, Heidelberg, 1993.
(18) Bevington, P. R. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill: New York, 1969.
(19) Press: W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flanner, B. P. *Numerical Recipes in Fortran 77. The Art of Scientific Computing*; Cambridge University Press: Cambridge, U.K., 1992.
(20) O'Connor, D. V.; Philips, D. *Time-Correlated Single Photon Counting*; Academic Press: London, 1984.
(21) Golub, G. H.; Pereyra, V. *SIAM J. Numer. Anal.* **1973**, *10*, 413−432.
(22) Golub, G.; Pereyra, V. *Inverse Problems* **2003**, *19*, R1−R26.
(23) Fisz, J. J. *Chem. Phys.* **1995**, *192*, 163−188.
(24) Charbonneau, P.; Knapp, B. *A user's Guide to PIKAIA 1.0*; NCAR Technical Note 418+IA; National Center for Atmospheric Research: Boulder, CO, 1996.
(25) Fisz, J. J.; Kołupajło, I. *Proc. SPIE Int. Soc. Opt. Eng.* **2005**, *5849*, 104−112.
(26) Boens, N.; van Dommelen, L.; Ameloot, M. *Biophys. Chem.* **1993**, *48*, 301−310.
(27) Birks, J. B. *Photophysics of Aromatic Molecules*; John Wiley and Sones: London, 1970.
(28) Meerts, W. L.; Schmitt, M.; Groenenboom, G. C. *Can. J. Chem.* **2004**, *82*, 804−819.
(29) Meerts, W. L.; Schmitt, M. *Phys. Scr.* **2006**, *73*, C47−C52.
(30) Meerts, W. L.; Schmitt, M. *Int. Rev. Phys. Chem* **2006**, *25*, 353−406.