ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Analytical Biochemistry 312 (2003) 57–65

ANALYTICAL BIOCHEMISTRY

www.elsevier.com/locate/yabio

Analysis of exponential data using a noniterative technique: application to surface plasmon experiments

Raimund J. Ober,[a,b] Jeffrey Caves,[c] and E. Sally Ward[b,c,*]

[a] *Center for Systems, Communications and Signal Processing, Eric Jonsson School of Electrical Engineering and Computer Science, University of Texas at Dallas, Richardson, TX 75083-0688, USA*

[b] *Cancer Immunobiology Center NB9.106, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390-8576, USA*

[c] *Center for Immunology, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390-8576, USA*

Received 8 July 2002

Abstract

The analysis of experimental data of exponential type plays a central role in many biophysical applications. We introduce a novel noniterative algorithm to analyze the association phase and dissociation phase of surface plasmon resonance experiments. It is shown that this algorithm can determine kinetic constants with a high level of accuracy in the presence of significant levels of noise. This algorithm should provide a valuable alternative to existing data analysis techniques.

Keywords: Surface plasmon resonance; Subspace algorithm; Parameter estimation

The analysis of empirical data is often a key aspect of an experiment. The advent of high-powered and relatively inexpensive computers has brought with it renewed interest in the use of more complex algorithmic tools that had earlier not been possible to implement. One characteristic of the data of many biophysical experiments is that the data are in effect of exponential type, i.e., that the data are the sum of data points that can be represented as an exponential function with appropriately chosen coefficient and exponent. Surface plasmon resonance experimental data, in the absence of mass transport, is the example that will be investigated here. Liquid-phase NMR (see, e.g., [1–3]), fluorescence life time (see, e.g., [4]), and sedimentation equilibrium (see, e.g., [5]) experiments also lead to this type of data.

Early approaches to the analysis of BIAcore dissociation data for simple 1:1 interactions relied on taking the logarithm of the data and on determining the dissociation constant through a simple linear regression analysis [6–8]. This approach is not very satisfactory for a number of reasons. For example, taking the logarithm destroys many basic statistical properties of the noise, which can lead to nonoptimal estimates. This approach can also be applied only to data resulting from 1:1 interactions with nonnegative data points. Therefore, use of a nonlinear least squares optimization routine to fit a suitable model to the data is now often advocated [6,9]. In fact, we [10] have shown that this approach leads to very good estimates in the sense that the variance of the parameter estimates is consistent with the Cramer Rao lower bound. One major disadvantage is that the optimization routine is iterative and therefore suffers from the standard convergence problems of gradient-based iterative schemes. First, initial conditions have to be given which requires another algorithm. Second, if the initial conditions are not close to the actual global minimum of the optimization problem the algorithm might converge to a local minimum which corresponds to an incorrect parameter estimate.

An ever greater demand for experimental throughput also puts higher demands on the capabilities of the data analysis software. With existing tools it is not too difficult to analyze individual sensorgrams since potential shortcomings in the data analysis algorithms can be overcome by user interaction with the software.

* Corresponding author. Fax: 1-214-648-1259.
E-mail address: sally.ward@UTSouthwestern.edu (E. Sally Ward).

0003-2697/02/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
PII: S0003-2697(02)00431-1

High-throughput experimentation does, however, require algorithmic tools that depend less on user input.

A recent advance in the development of data analysis algorithms has received a considerable amount of attention in the signal processing field. The so-called *subspace* algorithms (see, e.g., [11–15]) are noniterative algorithms that exhibit high-quality estimates. In this paper we investigate the subspace algorithm that addresses the data analysis problem encountered in the analysis of surface plasmon resonance data as acquired for example on a BIAcore instrument.

It is the expectation that this algorithm can become a routine technique to analyze surface plasmon resonance data.

Materials and methods

The experimental setup and data have been described earlier [10]. The main aspects are briefly summarized here.

Reagents

The mouse antibody (IgG1) was purified from culture supernatants of the 9E10 hybridoma [16] using protein A–Sepharose and standard methods. The myc peptide (EQKLISEEDLN) was synthesized in the peptide synthesis facility of the Howard Hughes Medical Institute and Department of Biochemistry, University of Texas Southwestern Medical Center.

Immobilization

The 9E10 antibody was coupled at various densities to CM5 chips using amine coupling. Relatively high coupling densities (ca. 5000–12,000 resonance units (RU)) were found to be necessary to obtain measurable signals with the myc peptide. However, to avoid unacceptable baseline drift these high coupling densities necessitated the use of extensive equilibration times prior to running experiments. For use as a control (reference) surface, flow cell 1 of each sensor chip was treated with the coupling chemistry using buffer without added ligand.

SPR experiments

Experiments were carried out using a BIAcore 2000 instrument (see [6]). Experiments were run using programed methods and the BIAcore control software. Myc peptide was used at a concentration of 38 μM in phosphate-buffered saline, pH 7.2, containing 0.01% Tween 20 (PBST). All experiments were run at 25 °C and injections of analyte were carried out using 240 μl and the kinject command. Dissociation phases were sufficiently long to allow complete dissociation of analyte from the flow cells, therefore avoiding the need to use regeneration injections. A flow rate of 20 μl/min was used in all experiments. For the present data 11 repetitions of the experiment were analyzed.

Data analysis

Data models

The standard equation that describes the result of a measurement in a flow cell for a 1:1 interaction is given by (see, e.g., [6])

$$\frac{\mathrm{d}R}{\mathrm{d}t}(t) = k_{\mathrm{on}}C(t)(R_{\max} - R(t)) - k_{\mathrm{off}}R(t),$$

where $R(t)$ is the measured signal in resonance units (RU), k_{on} is the association rate constant, k_{off} is the dissociation rate constant, $R_{\max}$ is the maximum analyte binding capacity in (RU), and $C(t)$ is the concentration of the analyte that flows over the chip.

During the association phase the analyte concentration is assumed to be constant, i.e., $C(t) =: C_0$ for $t_0 \leqslant t \leqslant t_1$. During the subsequent dissociation phase the injected analyte concentration is then set to zero; i.e., $C(t) = 0$ for $t \geqslant t_1$. Hence for the association phase the signal is given by

$$R_{\mathrm{a}}(t) = R_{\mathrm{eq}}(1 - e^{-tk_{\mathrm{obs}}}) \quad \text{for } 0 \leqslant t \leqslant t_1,$$

where $k_{\mathrm{obs}} := k_{\mathrm{on}}C_0 + k_{\mathrm{off}}$ and $R_{\mathrm{eq}} := (k_{\mathrm{on}}C_0R_{\max})/k_{\mathrm{obs}}$; for the dissociation phase it is given by

$$R_{\mathrm{d}}(t) = R_{\mathrm{a}}(t_1)e^{-(t-t_1)k_{\mathrm{off}}} \quad \text{for } t \geqslant t_1.$$

It is important to note for the subsequent development that both the association and the dissociation phase data are of exponential type; i.e., they are linear combinations of exponential functions. Here we also interpret the constant function as an exponential function with zero exponent. Specifically, if we set

$$A_{\mathrm{d}} := k_{\mathrm{off}}, \quad b_{\mathrm{d}} := R_{\mathrm{a}}(t_1), \quad \text{and} \quad c_{\mathrm{d}} := 1$$

then the dissociation data is easily represented as

$$R_{\mathrm{d}}(t) = c_{\mathrm{d}}e^{(t-t_1)A_{\mathrm{d}}}b_{\mathrm{d}} \quad \text{for } t \geqslant t_1.$$

Similarly, if we set

$$A_{\mathrm{a}} := \begin{pmatrix} 0 & 0 \\ 0 & -k_{\mathrm{obs}} \end{pmatrix}, \quad b_{\mathrm{a}} := \begin{pmatrix} R_{\mathrm{eq}} \\ -R_{\mathrm{eq}} \end{pmatrix}, \quad c_{\mathrm{a}} := \begin{pmatrix} 1 & 1 \end{pmatrix}$$

then the association phase data can be represented as

$$R_{\mathrm{a}}(t) = c_{\mathrm{a}}e^{tA_{\mathrm{a}}}b_{\mathrm{a}} \quad \text{for } t \geqslant 0,$$

where we used here the standard definition of the matrix exponential (see, e.g., [17]).

The measured data are of course not of the continuous time type described here but are sampled data with constant sampling intervals and are corrupted by noise, i.e., the data are given by

$$y(kT) = ce^{kTA_{\mathrm{c}}}b + w(k) = cA^kb + w(k) \quad \text{for } k = 0, 1, 2, \ldots, \tag{1}$$

for some vectors c, b, and matrix A_c, with $A = e^{TA_c}$, sampling interval T, and noise sequence $w(k)$ for $k = 0, 1, 2, \ldots$.

Subspace algorithm for exponential data

The specific subspace algorithm (see, e.g., [12,13]) that we will review here determines the A, b, and c in Eq. (1), given a data sequence $y(0), y(1), \ldots, y(N)$.

First, the original data is organized into a Hankel matrix

$$H_{\text{data}} = \begin{bmatrix} y(0) & y(1) & \cdots & \cdots & \cdots \\ y(1) & y(2) & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & y(N-2) & y(N-1) \\ \vdots & & & y(N-1) & y(N) \end{bmatrix}.$$

Second, a singular value decomposition $H_{\text{data}} = USV$ of the Hankel matrix is performed, where U and V are unitary matrices, i.e., $U^*U = I$ and $VV^* = I$, and S is a diagonal matrix with positive entries, the singular values, that are ordered with respect to magnitude, i.e., the largest singular value is the first diagonal entry, etc.

The subsequent steps depend on the number of components that make up the signal that is to be estimated. Let k be the number of exponential components that make up the data. In the situation discussed above, a 1:1 interaction, $k = 1$ for the dissociation data and $k = 2$ for the association data. We define reduced matrices of U, V, and S as follows. U_{reduced} is given by the first k columns of U, V_{reduced} is given by the first k rows of V, and S_{reduced} is given by the first k rows and columns of S.

The final estimates $\hat{A}$, $\hat{b}$, and $\hat{c}$ of A, b, and c are now obtained as follows. The estimate $\hat{c}$ is given by the first row of $U_{\text{reduced}}S_{\text{reduced}}^{1/2}$, the estimate $\hat{b}$ is given by the first column of $S_{\text{reduced}}^{1/2}V_{\text{reduced}}$, and $\hat{A}$ is given by

$$\hat{A} = U_{\text{reduced}}^{\uparrow} S_{\text{reduced}}^{1/2} (U_{\text{reduced}}^{\downarrow} S_{\text{reduced}}^{1/2})^{-\text{R}},$$

where $U_{\text{reduced}}^{\uparrow}$ is obtained from U_{reduced} by removing the first row of U_{reduced}, and $U_{\text{reduced}}^{\downarrow}$ is obtained from U_{reduced} by removing the last row of U_{reduced} and $\cdot^{-\text{R}}$ stands for a right inverse of $\cdot$, i.e. $\hat{\mathbf{A}}$ can be obtained by solving the associated system of linear equations.

As a last step the desired parameters have to be extracted from the estimated $\hat{A}, \hat{b}$, and $\hat{c}$. To this end let W be a diagonalizing transformation of $\hat{A}$; i.e., $A_{\text{D}} = W\hat{A}W^{-1}$ is a diagonal matrix. Also define $b_{\text{D}} = W\hat{b}$ and $c_{\text{D}} = \hat{c}W^{-1}$. Of course, for the dissociation data discussed above this step is not necessary since $\hat{A}$ is a scalar. In this case we can immediately set $A_{\text{D}} := \hat{A}$, $b_{\text{D}} := \hat{b}$, and $c_{\text{D}} := \hat{c}$.

For the dissociation data the dissociation constant k_{off} is estimated by

$$\hat{k}_{\text{off}} = -\frac{\log(A_{\text{D}})}{T},$$

where T is the sampling interval. The analogous step for the association data is slightly more delicate. For consistency of notation, we assume that the first diagonal entry of the 2×2 matrix A_{D} is closest to 1. If this is not the case, the diagonal entries of A_{D}, the entries of b_{D}, and the entries of c_{D} need to be transposed. Let now a_1 be the first and a_2 be the second diagonal entries of A_{D}. An estimate of $k_{\text{obs}} = k_{\text{on}}C_0 + k_{\text{off}}$ is given by

$$\hat{k}_{\text{obs}} = -\frac{\log(a_2)}{T},$$

and an estimate, $\hat{R}_{\text{eq}}$, of the equilibrium value R_{eq} is given by

$$\hat{R}_{\text{eq}} = b_{\text{D}}(1)c_{\text{D}}(1).$$

The analysis of data for other interaction models such as the parallel association and dissociation of two ligands is carried out analogously.

Data analysis based on logarithm of data

To analyze the data in the "classical" approach [7] the logarithm of all the dissociation data points is taken. For the association data the equilibrium value R_{eq} is estimated by averaging a number of data points for which the equilibrium value of the association phase has been attained. This equilibrium value is then added to each of the association data points after these data have been multiplied by -1. The logarithm of the resulting data points is taken.

The observed association constant k_{obs} and dissociation constant k_{off} are then easily estimated by solving a linear regression problem.

Parameter estimation by least squares minimization

The initial conditions for the gradient-based least squares algorithm were given by a subspace estimate. The gradient-based search algorithm to minimize the least squares modeling error was of the Marquardt–Levenberg type [18].

Simulation of BIAcore data

Data were simulated using the programming language Matlab [19]. Zero mean independent Gaussian noise was added to each data point. The standard deviation was chosen to be 0.25 RU in line with the measured noise level in our BIAcore 2000 instrument (see [10]), which is slightly lower than the noise standard deviation of 0.3 RU as reported in the instrument manual [6].

Data processing and software environment

All data were analyzed and processed using custom-written software in the programming language Matlab. The sensorgrams were all zero-adjusted and reference cell data were subtracted. For the analysis of the association

and dissociation phases suitable data segments were cut out from the sensorgrams. The software can be obtained by contacting the first author (ober@utdallas.edu).

Results

To evaluate the performance of the proposed subspace algorithm a data set with relatively low signal levels was chosen. Such data sets are typically more difficult to analyze since the signal-to-noise ratio is lower and potential artifacts are more pronounced than those in data sets with larger signal levels. Fig. 1 shows an overlay of the 11 sensorgrams for the myc peptide–9E10 antibody interaction [10,16].

All of the sensorgrams in this data set were analyzed using the methods presented above. Fig. 2 shows a representative association data segment analyzed using

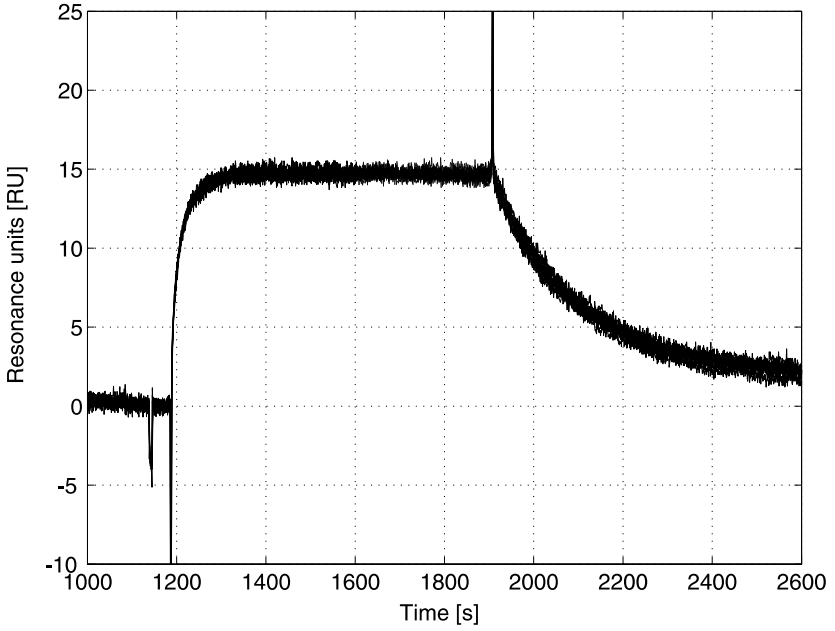

Fig. 1. An overlay of 11 sensorgrams of the myc peptide–9E10 antibody interaction [10] is shown. The measured sensorgrams are zero-adjusted and aligned along the time axis, and a background reference sensorgram is subtracted. Myc peptide was injected at a concentration of 38 μM and a flow rate of 20 μl/min.

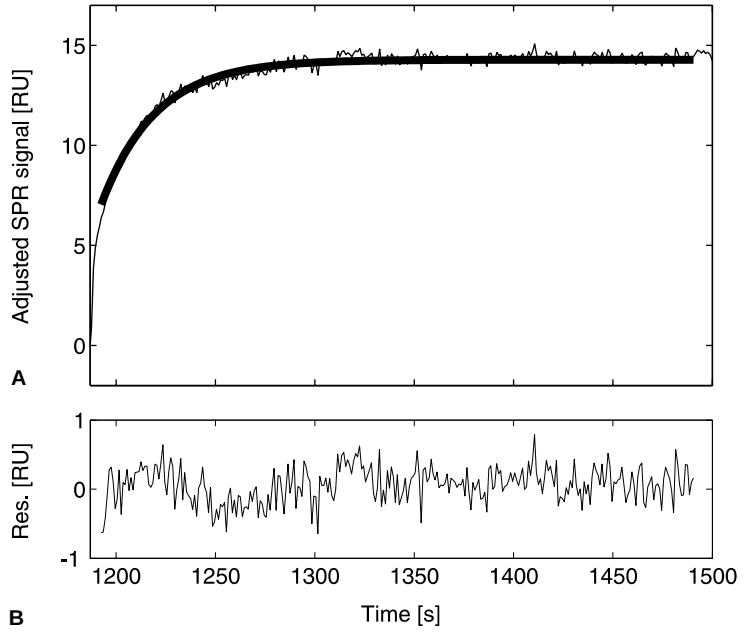

Fig. 2. (A) The association phase of one of the sensorgrams of the data presented in Fig. 1. The data fit based on the subspace method is shown in heavy solid line. (B) The residuals of the fit in A.

the subspace method. The plot of the residuals shows that the fit is good overall. A measure of the quality of the fits is that a subsequent adjustment of the parameters using a gradient-based optimization routine did not significantly reduce the mean square errors. This is shown in Figs. 3 and 4, where the values of the estimates are shown for the rate constants for the individual sensorgrams and the corresponding mean square errors. The results are summarized in Table 1. They show that use of the subspace-based method generally produces

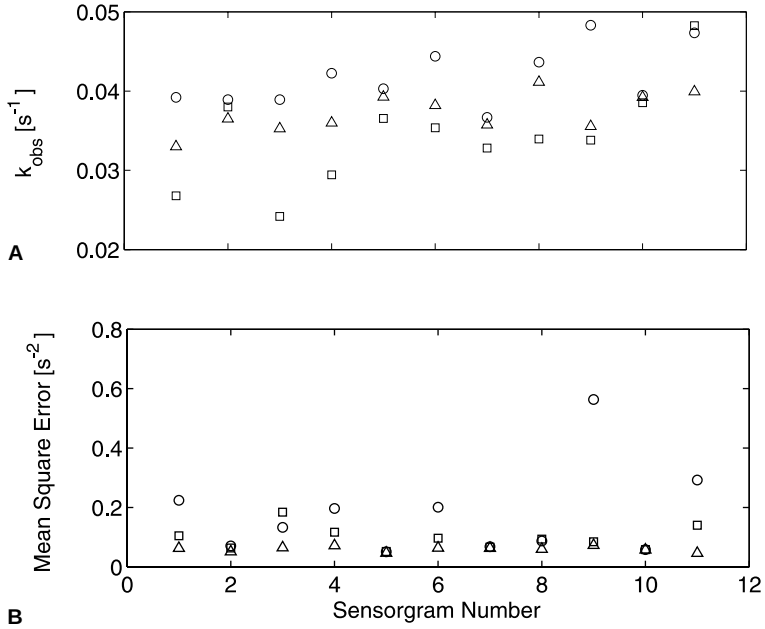

Fig. 3. (A) Estimates of $k_{obs} = k_{on}C_0 + k_{off}$ for the association phases of the sensorgrams in Fig. 1. The parameters from three estimation methods are shown: linear regression of the log of the data (□), subspace method (○), and optimization method using the results of the subspace method as initial conditions (△). (B) The mean square errors for the fits that yielded the estimates in A. The symbols correspond to those in A.

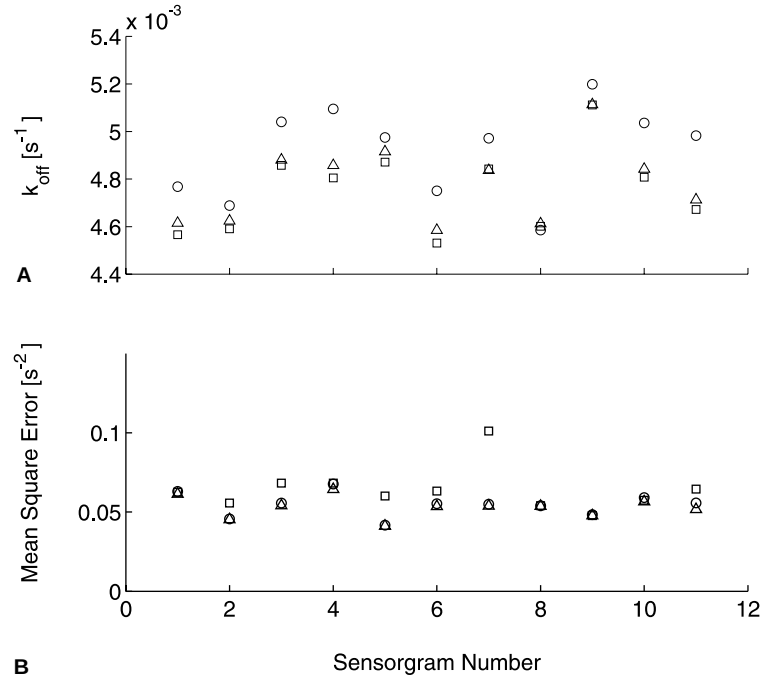

Fig. 4. A Estimates of the dissociation constant k_{off} for the dissociation phases of the sensorgrams in Fig. 1. The parameters from three estimation methods are shown: linear regression of the log of the data (□), subspace method (○), and optimization method (△). (B) The mean square errors for the fits that yielded the estimates in A. The symbols correspond to those in A.

Table 1
Myc–9E10 antibody interaction

Estimation method[a]	Mean k_{obs} (s^{-1})	Var k_{obs} (s^{-2})	Mean R_{eq} (RU)	Var R_{eq} (RU^2)	Average mean square error (RU^2)
Linear regression	0.034334	4.1767e−005	14.67	0.040272	0.096408
Subspace	0.04177	1.4059e−005	14.2638	0.099033	0.17685
Optimized	0.037254	6.0447e−006	14.642	0.032477	0.06018
Estimation method[b]	Mean k_d (s^{-1})	Var k_d (s^{-2})	Mean R_0 (RU)	Var R_0 (RU^2)	Average mean square error (RU^2)
Linear regression	0.0047508	3.0753e−008	13.7255	0.18459	0.063886
Subspace	0.0049176	3.6375e−008	13.801	0.14189	0.054625
Optimized	0.0047818	2.749e−008	13.801	0.13984	0.053033

[a] Results of estimates for the analysis of the association phase of the myc peptide–9E10 antibody interaction (Fig. 1). For each of the 11 association segments in this data set, the observed association constant k_{obs} and the equilibrium value R_{eq} are estimated with the linear regression method, the subspace method, and the optimized nonlinear least squares method using the estimates of the subspace model as initial conditions. For each estimation method, the mean and variance of the 11 k_{obs} and R_{eq} estimates are tabulated. In addition, the average mean square error of the fits is shown for each estimation method. The length of the association segment is varied to use the most suitable segment for each type of estimation method. In particular, for the linear regression approach only very short segments could be used. Otherwise, negative values in the difference between the data and the estimated equilibrium values would cause the logarithm operation to fail.

[b] Results of estimates for the analysis of the dissociation phase of the myc peptide–9E10 antibody interaction (Fig. 1). For each of the 11 dissociation segments in this data set, the dissociation constant k_d and the initial value R_0 are estimated with the linear regression method, the subspace method, and the optimized nonlinear least squares method using the estimates of the subspace model as initial conditions. For each estimation method, the mean and variance of the 11 k_d and R_0 estimates are tabulated. Again, the average mean square error of the fits is shown for each estimation method. The length of the dissociation segment is 70 data points in all cases.

good results that are often better than those obtained by use of the linear regression-based method. The optimization-based results consistently show the smallest mean square error. This is not surprising because the subspace-based estimates were used as initial conditions for the optimization algorithm, and the algorithm minimizes mean square error.

The linear regression-based method is applicable only for 1:1 interactions, whereas the subspace method can be used to analyze exponential data that arise from more complicated interactions. To demonstrate this, noisy data were simulated for a parallel association and dissociation interaction for two ligands (see Fig. 5 for a representative simulated sensorgram). Using the subspace-based method a data set of 20 sensorgrams was analyzed. The two association constants and dissociation constants, the corresponding coefficients, and the equilibrium value R_{eq} were esti-

Table 2
Parallel interaction with two ligands

	Parameter value	Subspace		Optimized	
		Mean	Variance	Mean	Variance
Dissociation phase					
k_{off1} (s^{-1})	0.1	0.10097	3.6117e−005	0.099882	3.137e−005
k_{off2} (s^{-1})	0.02	0.020217	1.3281e−006	0.020008	1.1633e−006
R_{01} (RU)	10	9.9401	0.15804	10.0008	0.13583
R_{02} (RU)	5	5.0723	0.189	4.9971	0.15925
Association phase					
k_{obs1} (s^{-1})	0.02	0.020641	3.701e−006	0.019696	1.7156e−006
k_{obs2} (s^{-1})	0.5	0.50848	0.0092195	0.5215	0.0031465
R_{eq1} (RU)	10	9.9182	0.39673	10.1084	0.12752
R_{eq2} (RU)	5	4.7914	0.11445	4.9759	0.032272
$R_{eq1} + R_{eq2}$ (RU)	15	14.7096	0.65333	15.0843	0.079072
R_{eq} (RU)	15	14.8721	0.52575	—	—

Results of the analysis of the simulated parallel association and dissociation interaction (Fig. 5) are shown. The parameter value column gives the values used to simulate the association and dissociation data. Independent Gaussian noise with zero mean and standard deviation of 0.25 RU has been added to the data points. The subspace method and optimized nonlinear least squares method (using the estimates of the subspace method as initial conditions) have been used to estimate the parameters. The simulation has been repeated 20 times with randomly generated noise. For each estimation technique, the means and variances of the various parameters estimated for the 20 sensorgrams are tabulated. Note that in this model there are two ways to estimate the equilibrium value R_{eq} of the association phase of the sensorgram. First, the coefficients R_{eq1} and R_{eq2} can be added. Second, R_{eq} can be estimated directly as was done in the analysis of the 1:1 interaction (see Materials and methods). The results of both approaches are given. They agree well, but not surprisingly the variance of the direct estimate of R_{eq} is slightly lower than that obtained as the sum of R_{eq1} and R_{eq2}. The average mean square error for the subspace models is 0.29313 RU^2 for association and 0.06369 RU^2 for dissociation. For the optimized models, the corresponding error measures are 0.07027 RU^2 for association and 0.06363 RU^2 for dissociation.

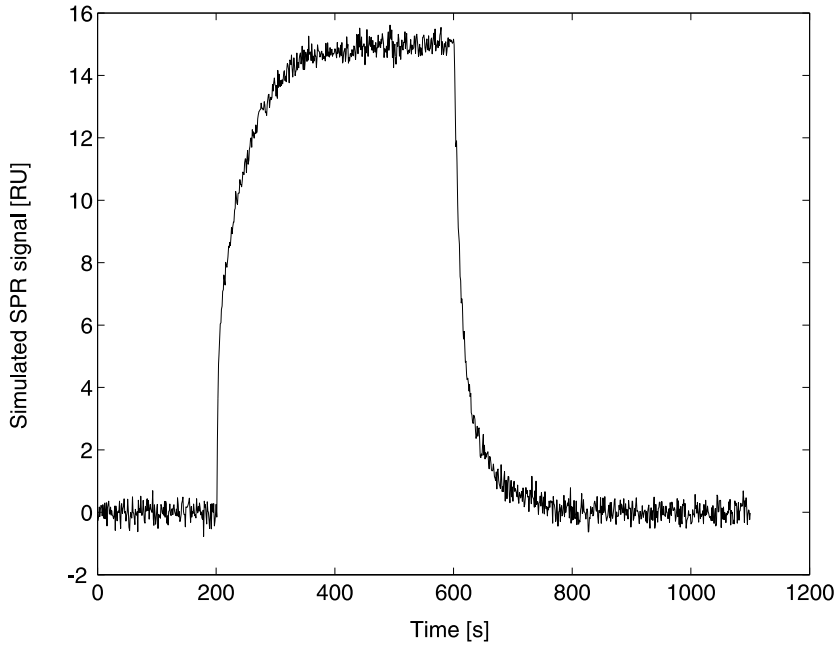

Fig. 5. A simulated sensorgram for a parallel association and dissociation interaction of two ligands, i.e., the sensorgram is given by

$$R(t) = \begin{cases} 0 & t < t_1 \\ R_{\mathrm{eq1}}(1 - e^{-(t-t_1)k_{\mathrm{obs1}}}) + R_{\mathrm{eq2}}(1 - e^{-(t-t_1)k_{\mathrm{obs2}}}) & t_1 \leqslant t < t_2 \\ R_{01}e^{-(t-t_2)k_{\mathrm{d1}}} + R_{02}e^{-(t-t_2)k_{\mathrm{d2}}} & t \geqslant t_2 \end{cases}$$

with $t_1 = 200\,\mathrm{s}$, $t_2 = 600\,\mathrm{s}$, $k_{\mathrm{obs1}} = 0.02\,\mathrm{s}^{-1}$, $k_{\mathrm{obs2}} = 0.5\,\mathrm{s}^{-1}$, $R_{\mathrm{eq1}} = 10$ RU, $R_{\mathrm{eq2}} = 5$ RU, $k_{\mathrm{off1}} = 0.1\,\mathrm{s}^{-1}$, $k_{\mathrm{off2}} = 0.02\,\mathrm{s}^{-1}$, $R_{01} = 10$ RU, $R_{02} = 5$ RU. Thus, the association phase is simulated for 400 s and subsequent dissociation is simulated for 500 s. Gaussian noise with a standard deviation of 0.25 RU is added to the data points.

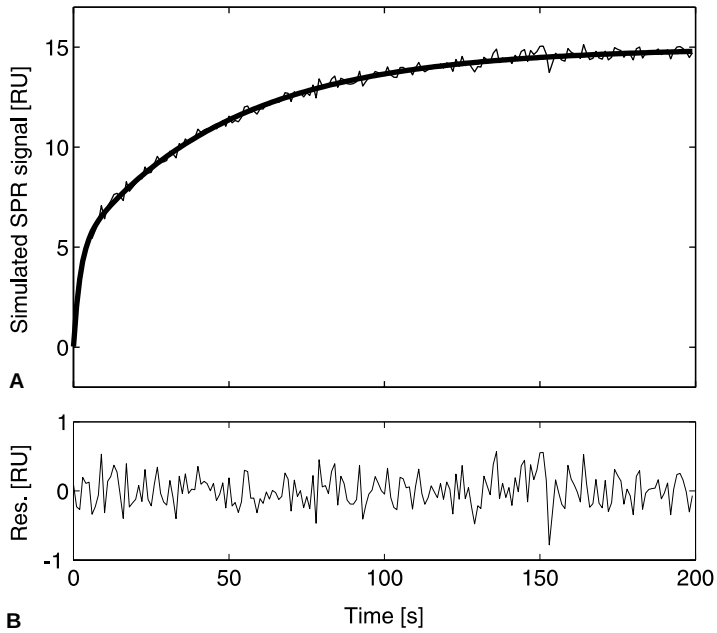

Fig. 6. (A) The association phase of the simulated sensorgram described in Fig. 5 with data fit based on the subspace method (heavy solid line). (B) The residuals of the fit in (A).

mated. Fig. 6 shows a representative fit of the association phase of a sensorgram. The results are summarized in Table 2. The estimates using the subspace-based method generally show only a small difference compared with the actual values of the parameters. These estimates were further improved by

using them as initial conditions for a gradient-based optimization algorithm.

Discussion

Exponential data arises in many biophysical applications. Here our particular interest is data arising from surface plasmon resonance experiments. However, the approach and results are easily applied to many other areas of biophysical data analysis.

The analysis of such exponential data is not without problems. If no measurement noise were present a number of more or less elementary schemes would easily lead to a solution. It is, however, the presence of noise that complicates the analysis. In most cases the parameters of interest are decay rates. This means that the parameters to be estimated are nonlinear in the data. Currently two data analysis methods are primarily used. In one method, the logarithm of the data points is taken [7]. This translates the data analysis problem into a linear regression problem if the data are made up of only one exponential component. In the other standard method, gradient-based optimization algorithms are used to fit a model to the data [6,9]. Both algorithms have their advantages and disadvantages. The first method is restricted to the analysis of 1:1 interactions. In addition, the logarithm introduces nonlinear distortions in the noise and changes the noise characteristics in a nonuniform fashion. This makes a statistical analysis of the estimation procedure difficult and can lead to biased estimates. Furthermore, when low signal levels are encountered, measurement of noise might lead to negative data points for which the logarithm cannot be taken. The gradient-based optimization methods on the other hand often have statistically optimal or near optimal properties (see, e.g., [10]). Their disadvantages lie in the fact that convergence is not guaranteed. Depending on the initial conditions of the search the algorithm could converge to a local minimum rather than a global one. This highlights another potential problem with iterative algorithms, i.e., that good initial conditions are needed. An additional algorithm is needed to determine these initial conditions.

In the signal-processing literature so-called subspace algorithms have received a significant amount of attention (see, e.g., [11–15]). Here we have analyzed simulated and experimental surface plasmon resonance data acquired on a BIAcore 2000 instrument using a particular subspace algorithm for exponential type data. Subspace algorithms have the advantage that they are noniterative and therefore do not require initial conditions and do not have the convergence problem of iterative algorithms. In contrast to the linear regression-based algorithm that relies on taking the logarithm of the data points, subspace algorithms are not restricted to data that arise from 1:1 interactions and can easily deal with negative data points. We have adapted the existing subspace algorithm to analyze the present data and evaluated its performance using both experimental and simulated data.

We have examined the use of this algorithm to provide estimates for association and dissociation data. We found that the estimates are very accurate. In fact, if the estimates are used as starting points for a subsequent gradient-based optimization routine the quality of the parameter estimates does not improve significantly through the further optimization step.

An examination of the performance of the subspace algorithm applied to simulated data for a parallel association and dissociation interaction of two ligands shows that the algorithm also performs well in this more complicated situation. Such an interaction cannot be adequately analyzed with the standard method which relies on turning the problem into a linear regression problem by taking logarithms of the data.

The use of this subspace algorithm is, however, also not without potential pitfalls. The algorithm can easily handle very high noise levels, but care should be taken with data sets that have outliers or other artifacts. These should be removed; the resulting gaps can be filled with values interpolated from neighboring data points.

A central step in the subspace method is a singular value decomposition. This is typically a computationally expensive operation. However, with current computers this is not a problem for the typical lengths of data sets that are being analyzed. In fact, in our experience the time taken to analyze a data set using the subspace method was significantly shorter than the time taken by the gradient-based method.

This algorithm, like most others, will not produce reliable estimates if the data does not comply with the model. For example, in the estimation of dissociation data for a 1:1 interaction a misleading estimate could be expected when the data have a nonzero offset that is not accounted for in the model. To account for such an offset is in principle straightforward and would be a minor modification of the procedure used to estimate the equilibrium value R_{eq} in the analysis of the association data.

The newly proposed method is a complement to existing methods and provides an important tool for the analysis of surface plasmon resonance experiments of exponential type. The availability of noniterative algorithms is of particular importance for automated data analysis problems.

Acknowledgment

This research was supported in part by a grant from the National Institutes of Health (NIH GM58538).

References

[1] G. Bodenhausen, R.R. Ernst, A. Wokaun, Principles of Nuclear Magnetic Resonance in One and Two Dimensions, Oxford Science Publications, 1987.

[2] R.J. Ober, E.S. Ward, Correcting for phase distortion of NMR spectra analyzed using singular value decomposition of Hankel matrices, J. Magn. Reson. Ser. A 114 (1995) 120–123.

[3] R.J. Ober, E.S. Ward, System theoretic formulation of NMR experiments, J. Math. Chem. 20 (1996) 47–65.

[4] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, second ed., Kluwer Academic/Plenum Publishers, New York, 1999.

[5] P. Schuck, Sedimentation equilibrium analysis of interface optical data by systematic noice decomposition, Anal. Biochem. 272 (2) (1999) 199–208.

[6] Biotechnology Handbook, Pharmacia Biosensor AB, Uppsala, 1994.

[7] A. Michaelsson, R. Karlsson, L. Mattsson, Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system, J. Immunol. Methods 145 (1991) 229–240.

[8] M. Ohlin, B. Jansson, A.-C. Malmborg, A. Michaëlsson, C.A.K. Borrebaeck, Real time analysis of antibody-antigen reaction kinetics, Scand. J. Immunol. 35 (1992) 643–650.

[9] K.K. Soneson, D.J. O'Shannessey, M. Brigham-Burke, P. Hensley, Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis method, Anal. Biochem. 212 (1993) 457–468.

[10] R.J. Ober, E.S. Ward, The influence of signal noise on the accuracy of kinetic constants measured by surface plasmon resonance experiments, Anal. Biochem. 273 (1999) 49–59.

[11] P.V. Overschee, B. De Moor, Subspace Identification for Linear Systems: Theory, Implementation, Applications, Kluwer Academic Publishers, Dordrecht, 1996.

[12] H.P. Zeiger, A.J. McEwen, Approximate linear realization of given dimension via Ho's algorithm, IEEE Trans. Automat. Contr. AC-19 (1974) 153.

[13] S.Y. Kung, A new low-order approximation algorithm via singular value decomposition, in: Proc. 12th Asilomar Conf. Circuits, Syst. Computers, 1978.

[14] W.E. Larimore, Canonical variate analysis in identification, filtering and adaptive control, in: Proc. 29th IEEE Conf. Decision Contr., HI, 1990.

[15] W.E. Larimore, System identification, reduced-order filtering and modeling via canonical variate analysis, in: Proc. Am. Contr. Conf., San Francisco, CA, 1983.

[16] G. Ramsay, J.M. Bishop, G.I. Evan, G.K. Lewis, Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product, Mol. Cell. Biol. 5 (1985) 3610–3616.

[17] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.

[18] M.A. Branch, T. Coleman, A. Grace, Optimization Toolbox For Use With Matlab, The Mathworks, Natick, 1999.

[19] MATLAB: The Language of Technical Computing, The Mathworks, Natick, 2000.