An exhaustive search approach for chemical kinetics experimental data fitting, rate constants optimization and confidence interval estimation

Audrius Laurynėnas, Juozas Kulys

Vilnius University Institute of Biochemistry
Mokslininkų str. 12, LT-08662 Vilnius, Lithuania
audrius.laurynenas@gmail.com

Abstract. Finding reaction rate constants in a complex chemical mechanism is a complicated problem. These complications can be attributed mainly to the nonlinearity of such systems, where the least squares optimization procedure fail when experimental data are being fitted. Here we analyze two cases: analytical and numerical. From analytical case possible issues of optimization are found and methods to cope with them are derived. The complex case shows how an exhaustive search fitting approach works. Our results indicate that certain properties of systems can be found only by running grid search on entire rate constant set and minima found have practical value only if the properties of entire fitting score surface are known.

Keywords: chemical kinetics, rate constant, fitting, Powell algorithm, exhaustive search.

1 Introduction

The task of the reaction rate constants determination from the experimental data is a routine procedure within the field of chemical and biochemical research. The question how a reaction mechanism is affected by varying reaction rate constants, and how to relate the experimental data with a certain reaction mechanism is not trivial. There are many known practical frameworks and applications to fit experimental data [2, 5].

Traditionally, chemical rate constants of a reaction mechanism are determined from the initial rate or from an integral equation fitting (in simple cases, when an integral equation is known). Alternatively, it can be done by direct fitting of the rate constants using the least squares method, which involves numerical evaluation of the ordinary differential equations those describe the reaction mechanism.

Possible choices for least square fitting are KinFitSim, which allows to fit more than one curve at the same time, but all of those curves must come from observations of different reaction constituents of same reaction with same initial concentrations [16]. KINFIT II is designed to analyze ligand binding data with the possibility to use data of more than one experiment [14] and DynaFit, which can fit data of a single experiment at

a time [8]. All these methods suffer from some drawbacks. Initial rate approximation method (initial reaction rate dependence on initial reagents concentrations) uses only minor experimental curve information to determine one or two reaction rate constants (or alternatively – Michaelis and reaction rate constants). This approach requires experimental design effort in order to be able to calculate rate constants for limiting mechanism step from steady state approximation [2].

On the other hand, the methods that employ least squares optimization can use (in principle) all the information in experimental data about the reaction rate constants (due to existence and uniqueness of differential equations system solutions [3]), but typically suffer from stalling somewhere if the initial rate constants are too far from the actual parameters' set. Traditionally, the user of those fitting programs is expected to provide good initial constants set in order to get a good fit after optimization. There is little research done on what causes this ill behavior. This article analyzes the main causes that make the fitting procedure fail and offers a heuristic method to provide good initial rate constants.

2 Methods

Most fitting problems start with well known least squares method, which is formulated as

$$\text{least squares score} \stackrel{\text{def}}{=} \sum_{i=1}^{n} (x_{i,\text{exp}} - x_{i,\text{calc}})^2, \tag{1}$$

where $x_{i,\text{exp}}$ designates experimental data point, $x_{i,\text{calc}}$ – calculated data point, n – number of experimental points. This formulation is adequate only for datasets, which are not affected by data array length or initial experiment conditions. Such an example would be a single experimental curve.

When the least squares method is used on a small number of datasets from experiments with different initial concentrations and different number of data points, the formula is requires a simple modification to make fitting score independent on the number of data points. Coefficient $1/n$ in the formula below transforms the least squares sum to statistical data set property variance as defined in [11]:

$$\sigma^2 \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} (x_{i,\text{exp}} - x_{i,\text{calc}})^2. \tag{2}$$

Variance can be used to optimize one experiment data in same fashion as plain least squares method (see Eq. (1)). In case of more than one experimental dataset, we define fitting score as sum of experimental datasets variances

$$\text{fitting score} = \sum_{i=1}^{m} \sigma_i^2, \tag{3}$$

where m is number of experimental curves with different initial concentrations. This definition is valid only in cases when noise level is approximately constant at range of measured experimental values.

All numerical integrations were performed with LSODA numerical integration package [12] integrated into PYTHON SCIPY library. LSODA automatically selects Adams method for non-stiff equations and Gear method for stiff equations. All equations prior passing to LSODA were dedimesionalized due to stability and robustness reasons. The dimension of the rate constant is s^{-1} and $L\,mol^{-1}s^{-1}$ for the first and the second order reaction, respectively. The dimension of the concentration is $mol\,L^{-1}$. The equilibrium constant is dimensionless.

Final optimization of fitting score was performed with Powell optimization algorithm [13], which is also integrated in the same SCIPY library. Integrations and optimizations performed with numerical tolerances set to 10^{-8}. All constants optimizations were performed in exponential form with replacement rule $k \to 10^{k'}$.

3 Exponential case

Actually, there is no interest in complicated fitting of simple first order reaction

$$A \xrightarrow{k} B \frac{\mathrm{d}A}{\mathrm{d}t} = -kA, \qquad \frac{\mathrm{d}B}{\mathrm{d}t} = kB. \tag{4}$$
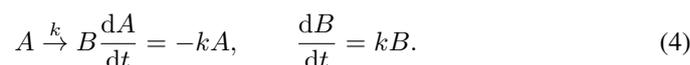
This mechanism has clear and simple solution $A(t) = A_0 e^{-kt}$. All experimental data can be fitted in this formula by obvious logarithmization. But this system has a fundamental virtue – our solutions of interest can be derived analytically.

Starting with the least squares approach and assuming experimentally measured concentration dependence on time for reagent A, which the initial concentration is 1, rate constant is k and the parameter to optimize – r, the fit score metric in least squares framework (see Eq. (1)) can be written as

$$f_{\text{fitting score}} = \int_0^T \left(e^{-kt} - e^{-rt}\right)^2 \mathrm{d}t. \tag{5}$$

Integration boundaries in Eq. (5) are replaced assuming that concentration of A is measured until it drops bellow $1/100$ of initial A ($T = \ln 100/k$). After integration, the fitting score is an analytical function with two variables k and r

$$\begin{aligned}
f_{\text{fitting score}} &= \int_0^{\ln 100/k} \left(e^{-kt} - e^{-rt}\right)^2 \mathrm{d}t \\
&= \frac{1}{2}\left(\frac{1 - 100e^{-2}}{k} + \frac{1 - 100e^{-2r/k}}{r} + \frac{4(-1 + 100e^{-(k+r)/k})}{k+r}\right),
\end{aligned} \tag{6}$$

where k is a constant to fit and r a is fitting parameter.

Exponential nature of the solution implies the use of logarithmic scale for parameters. Such simple transformation contracts curve parameters domain from range $1 \div 10^8$ to
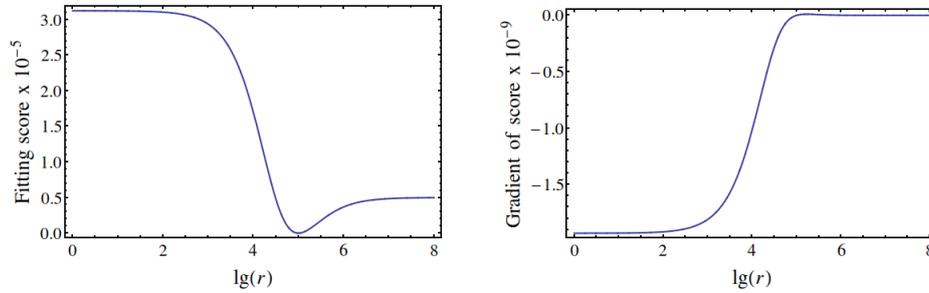
Fig. 1. Fitting score deppendence on fitting parameter r (in logarithmic scale) with $k = 10^5$.

range $0 \div 8$. Such a modification offers several advantages – ability to maintain the sign of optimized constants naturally and smoothly in the optimization procedure; also enhances the numerical stability of optimization (rises gradient values).

The solution of Eq. (6) and its derivative are plotted against r with arbitrary selected rate constant $k = 10^5$ in Fig. 1. Simple replacement $r = k$ yields 0. Graph of fitting score with $k = 10^5$ clearly shows minimum at the trial parameter. There is a more mathematically rigorous solution, which can be found by taking limit $T \to \infty$ (meaning infinitely large observation time):

$$\lim_{T \to \infty} \frac{1}{2} \left(\frac{4(e^{T(-k-r)} - 1)}{k + r} + \frac{1 - e^{-2kT}}{k} + \frac{1 - e^{-2rT}}{r} \right) = \frac{(k - r)^2}{2kr(k + r)}. \quad (7)$$

Taking derivative and equaling it to zero:

$$-\frac{(k - r)^2}{2kr(k + r)^2} + \frac{k - r}{kr(k + r)} - \frac{(k - r)^2}{2k^2 r(k + r)} = 0. \quad (8)$$

There are two solutions: $r_1 = -3k$ and $r_2 = k$. The first one can be ignored as physically meaningless.

At ranges of chemical interest this fit score function is continuous and has only one minimum. The most interesting feature is its flatness at sufficiently large distance from minima. This flatness is clearly responsible on small changes of gradient and it is main reason, why an optimization algorithm will fail at certain distance from actual solution – it just stalls on the flat surface with no directions where to move further. Practically, in this case, optimization problem is reduced to Easom function optimization [9]. The second cause of optimization failures is numerical errors. One can evaluate a gradient of fitting score in two points (with $k = 10^5$), let us say, $r_1 = 10$, $r_2 = 10^2$ and will get gradient values respectively $f'(r_1) = -3.8614 \cdot 10^{-9}$ and $f'(r_2) = -3.83865 \cdot 10^{-9}$. Differences of these values are on the edge of numerical precision.

This example resolves fitting question of system, which by itself is not complicated, and allows easy determination of constants from experimental data. But the main reason clearly seen from it is that the least square optimization for more complex systems can fail.

As the main reason of fitting failures could be the flat optimization surface of fitting score, a better method should be used to estimate the initial parameter for fitting. While those parameters can come from experiments or from intuition, we expect that rate constants could be more rigorously found by this procedure:
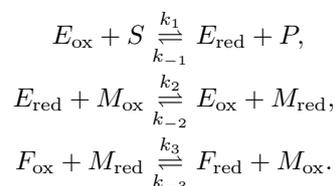
- Evaluation of fitting score on logarithmic rate constants grid with some arbitrary selected step size (an exhaustive search approach – well known general method for solving various problems).
- Selection of some set of grid points with best fitting score as initial guesses for optimization procedure.
- Optimization of selected initial guesses.
- Analysis of possible fitted rate constants values.

This approach is used in the next case.

4 Complex case

There is a vast range for selection of a test example. But we have a research interest in mediator driven enzymatic reactions. Due to limited enzymatic substrate specificity, it is impossible to use some enzymes directly, particularly, oxidoreductases to drive commercially and scientifically interesting reactions. However, this obstacle can be overcome with mediators – small molecules that can perform an intermediate function between enzyme and substrate of interest [10].

One of simplest cases of mediator driven enzymatic reactions can be represented as a system of three coupled equations (simplified mechanism from [7])

$$E_{\text{ox}} + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E_{\text{red}} + P,$$

$$E_{\text{red}} + M_{\text{ox}} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E_{\text{ox}} + M_{\text{red}},$$

$$F_{\text{ox}} + M_{\text{red}} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} F_{\text{red}} + M_{\text{ox}}.$$

In this scheme, we denote enzyme forms as E_{ox} and E_{red}, substrate for enzyme regeneration as S and its product as P. Substrate of interest and mediator as F and M, respectively. This case have no complications except reversibility in the third reaction. For the sake of simplicity, let us assume that the first two enzymatic reactions are bimolecular and nonreversible. These assumptions lead to the system of ordinary differential equations

$$\frac{\mathrm{d}[E_{\text{ox}}]}{\mathrm{d}t} = -k_1[E_{\text{ox}}][S] + k_2[E_{\text{red}}][M_{\text{ox}}],$$

$$\frac{\mathrm{d}[E_{\text{red}}]}{\mathrm{d}t} = k_1[E_{\text{ox}}][S] - k_2[E_{\text{red}}[M_{\text{ox}}],$$

$$\frac{\mathrm{d}[F_{\text{ox}}]}{\mathrm{d}t} = k_{-3}[F_{\text{red}}][M_{\text{ox}}] - k_3[F_{\text{ox}}][M_{\text{red}}],$$

$$\frac{\mathrm{d}[F_{\text{red}}]}{\mathrm{d}t} = -k_{-3}[F_{\text{red}}][M_{\text{ox}}] + k_3[F_{\text{ox}}][M_{\text{red}}],$$

$$\frac{\mathrm{d}[M_{\mathrm{ox}}]}{\mathrm{d}t} = -k_2[E_{\mathrm{red}}][M_{\mathrm{ox}}] + k_3[F_{\mathrm{ox}}][M_{\mathrm{red}}] - k_{-3}[F_{\mathrm{red}}][M_{\mathrm{ox}}],$$

$$\frac{\mathrm{d}[M_{\mathrm{red}}]}{\mathrm{d}t} = k_2[E_{\mathrm{red}}][M_{\mathrm{ox}}] - k_3[F_{\mathrm{ox}}][M_{\mathrm{red}}] + k_{-3}[F_{\mathrm{red}}][M_{\mathrm{ox}}],$$

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_1[E_{\mathrm{ox}}][S] \qquad \frac{\mathrm{d}[S]}{\mathrm{d}t} = -k_1[E_{\mathrm{ox}}][S].$$

All reagent concentrations in experiment could be known. In such case the problem reduces to fitting experimental data to four constants: k_1, k_2, k_{-3}, k_3. This system of nonlinear ordinary differential equations have no analytical solution. But this system can be integrated numerically and could be evaluated fitting score for some set of constants and experimental data curve.

It is necessary to begin analysis with global minimum search. As there is no analytical solution, some parameter set $\{k_1,\ k_2,\ k_{-3},\ k_3\}$ should be selected and numerically checked fitting score dependency on parameter variation. Let us fix our parameters $k_1 = 10^5$, $k_2 = 10^8$, $k_{-3} = 10^5$, $k_3 = 10^6$, which are quite realistic and also have some illustrative value. Initial concentrations can be selected also fairly realistically: $E_{\mathrm{ox}} = 10^{-9}$, $M_{\mathrm{ox}} = 10^{-6}$, $F_{\mathrm{ox}} = 10^{-4}$, $S = 10^{-3}$. Other concentrations set to zero. It is assumed that in reaction, F_{ox} concentration is measured (parameters were selected in analogy to [6, 7]).

Rate constants search grid is set in logarithmical form of size $9 \times 9 \times 9 \times 9$. There is a little interest to inspect bimolecular rate constants larger than 10^8 (or possibly 10^{10} – it deppends on predicted diffusion controled reaction rate constants limit [15]) and little use of reactions with bimolecular rate constants smaller than 10^0. Here test curve for reagent F_{ox} is simulated and is fitting score for entire system on grid evaluated.

As expected from simple case, fitting score for this complex mechanism exhibits same features – flatness of fitting score surface at large distance from simulation constants set and one global minimum of fitting score (Fig. 2). There are some doubts about practicality of such evaluation on constants grid and possibly quite large time and computation resources needed for this task. But actually, evaluation on grid of size $9 \times 9 \times 9 \times 9$ can be performed on Intel Core i3 single thread in less than 10 minutes (three data sets with 2000 data points) and on grid of $17 \times 17 \times 17 \times 17$ – in less than hour with Python script (with conventional program – possibly several order faster).

Of course, this initial set of parameters is quite idealistic, because evaluation grid by itself has point, in which fitting score is minimal. In real world, such perfect numbers as integer power of ten are rarely found. And of course, grid of constants could not be made infinitely small as in this case, computation time simply rises in order of fourth power. However there is no need to make grid finer. It could be simply selected some set of points with minimal fitting score and used as initial points for further minimization with some optimization algorithm. There is a hope that some points in this set will be quite close to global minimum and will converge to it.

For further testing of our procedure described at the end of simple case analysis, we select few sets of possible rate constants (rate constants sets and initial concentrations
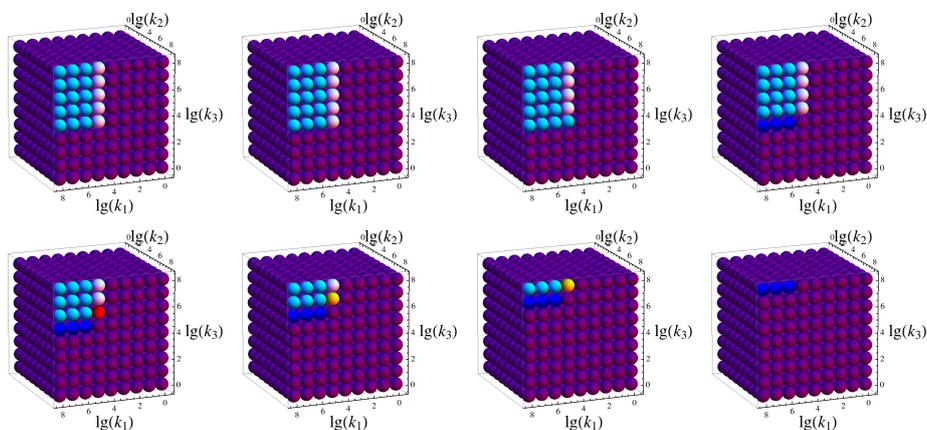
Fig. 2. Fitting score evaluation on constants grid, from left top to rigth bottom $k_{-3} = \{10^0, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8\}$. Purple – bad fit, red – perfect fit. (Online version in color.)
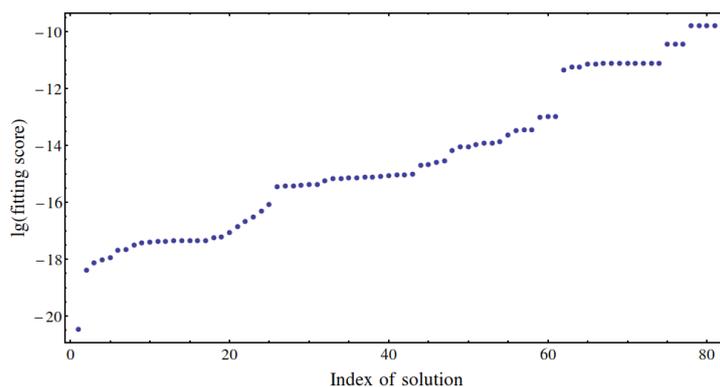

Fig. 3. Sorted fitting scores of optimized 81 initial grid search points (in logarithmic scale, unbiased case #1).

in Table 1). As procedure depends on number of best solutions after grid search, we found it useful to select 3^p (here p – number of constants to optimize, in this particular case, $3^4 = 81$) solutions for minimization with Powell algorithm. As was expected, initial grid points set after optimization does not converge to one solution. Actually, we are left with optimized solution set, where fitting score can vary in several orders of magnitude as showed in Fig. 3 (same behavior in other cases). It is worth to note that optimization procedure had terminated in each case without any flaws.

In Table 1, simulated rate constants and results of fitting on one kinetic curve are presented. All six cases behave in the same manner as the first one. Fitting score greatly varies from solution to solution and only in two cases, #4 and #6, best solution reasonably agrees with simulation parameters. These differences do not improve on finer grids (even worsen). Main reason of this behavior is possibly simple – one curve is not enough to fit four rate constants.

Table 1. Fitting constants of one kinetic curve to system of differential equations (coarse $9 \times 9 \times 9 \times 9$ grid). Initial concentrations $E_{ox} = 10^{-9}$, $M_{ox} = 10^{-6}$, $F_{ox} = 10^{-4}$, $S = 10^{-3}$.

Case	Simulation constants				Worst solution	Best solution
	$\lg(k_1)$	$\lg(k_2)$	$\lg(k_3)$	$\lg(k_{-3})$		
#1	6.3	7.3	6.3	5.3	fscore $= 1.63 \cdot 10^{-10}$	fscore $= 3.35 \cdot 10^{-21}$
					8.00 8.00 7.23 8.00	6.30 7.30 8.00 7.00
#2	5.3	7.3	6.3	5.3	fscore $= 1.70 \cdot 10^{-10}$	fscore $= 6.81 \cdot 10^{-17}$
					8.00 8.00 7.18 8.00	5.52 7.28 8.00 6.99
#3	5.3	6.3	6.3	5.3	fscore $= 2.22 \cdot 10^{-10}$	fscore $= 2.11 \cdot 10^{-19}$
					7.75 8.00 6.11 8.00	5.32 6.30 6.77 5.77
#4	5.3	6.3	5.3	5.3	fscore $= 4.04 \cdot 10^{-11}$	fscore $= 5.37 \cdot 10^{-21}$
					8.00 6.08 8.00 0.32	5.17 6.30 5.26 5.26
#5	5.3	6.3	5.3	4.3	fscore $= 7.72 \cdot 10^{-13}$	fscore $= 3.19 \cdot 10^{-23}$
					6.29 8.00 3.78 7.04	6.57 6.30 5.30 4.30
#6	5.0	8.0	7.0	7.0	fscore $= 3.41 \cdot 10^{-11}$	fscore $= 9.24 \cdot 10^{-21}$
					8.00 8.00 6.45 6.81	5.00 8.00 7.00 7.00

Table 2. Fitting constants of three kinetic curves to system of differential equations (coarse $9 \times 9 \times 9 \times 9$ grid). Initial concentrations $E_{ox} = 10^{-9}$, $M_{ox} = 10^{-(4 \div 6)}$, $F_{ox} = 10^{-4}$, $S = 10^{-3}$.

Case	Simulation constants				Worst solution	Best solution
	$\lg(k_1)$	$\lg(k_2)$	$\lg(k_3)$	$\lg(k_{-3})$		
#1	6.3	7.3	6.3	5.3	fscore $= 2.86 \cdot 10^{-10}$	fscore $= 1.10 \cdot 10^{-15}$
					6.75 8.00 5.54 6.30	6.30 7.30 6.43 5.43
#2	5.3	7.3	6.3	5.3	fscore $= 3.23 \cdot 10^{-10}$	fscore $= 1.48 \cdot 10^{-19}$
					5.60 8.00 7.17 8.00	5.30 7.30 6.30 5.30
#3	5.3	6.3	6.3	5.3	fscore $= 4.19 \cdot 10^{-10}$	fscore $= 2.11 \cdot 10^{-19}$
					5.64 8.00 6.12 8.00	5.30 6.30 6.24 5.24
#4	5.3	6.3	5.3	5.3	fscore $= 1.72 \cdot 10^{-10}$	fscore $= 4.14 \cdot 10^{-20}$
					4.69 6.09 6.86 0.00	5.30 6.30 5.35 5.35
#5	5.3	6.3	5.3	4.3	fscore $= 2.71 \cdot 10^{-11}$	fscore $= 2.87 \cdot 10^{-20}$
					5.63 8.00 1.42 2.74	5.30 6.30 5.24 4.24
#6	5.0	8.0	7.0	7.0	fscore $= 4.08 \cdot 10^{-10}$	fscore $= 5.29 \cdot 10^{-20}$
					4.78 8.00 8.00 2.24	5.00 8.00 7.00 7.00

This hypothesis can be tested with simultaneous fitting of three simulated experimental curves with different initial M_{ox} concentrations $- 10^{-4}, 10^{-5}, 10^{-6}$, as experimentally there are little difficulties to vary concentrations in reaction. The results of three curves fitting are summarized in Table 2. There is practically perfect agreement between rate constants used in simulations and rate constants, recovered after optimization.

As mentioned earlier, the set of solutions consists of 81 combinations of optimized rate constants. Even though perfect agreement between simulated and optimized parameters exist in this particular case with particular initial conditions, it is worth to analyze other solutions more closely. For example, we can take case #1, in which rate constants in one curve fitting vary greatly and are hard to interpret. 54 solutions of three curve fitting case shows stable values of k_1 and k_2. Rate constants k_3 and k_{-3} vary much more – from $10^{3.4}$ to 10^7 (k_{-3}), but their change is not independent – logarithm of constants differs by approximately one. This relation indicates two interesting and valuable findings about
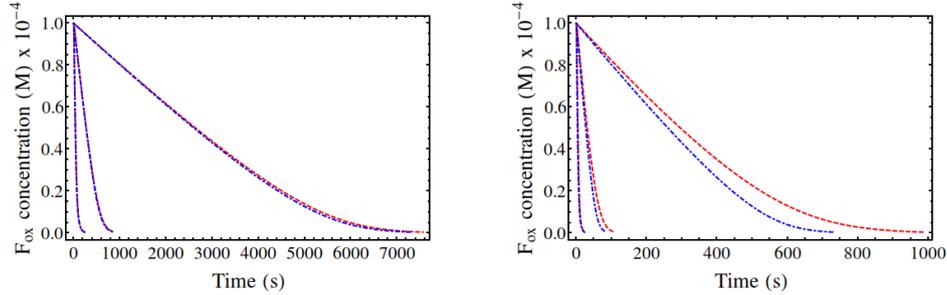
Fig. 4. Two sets of curves simulated with M_{ox} concentrations 10^{-6}, 10^{-5}, 10^{-4}. Red dashed – $k_1 = 10^{5.3}$, $k_2 = 10^{6.3}$, $k_3 = 10^{4.3}$, $k_{-3} = 10^{3.3}$; blue dashed – $k_1 = 10^{5.30}$, $k_2 = 10^{6.30}$, $k_3 = 10^8$, $k_{-3} = 10^7$. Left simulation $[E_{ox}] = 10^{-9}$, right – $[E_{ox}] = 10^{-8}$. (Online version in color.)

this particular mechanism – the third reaction is in equilibrium and overall reaction rate is not affected by absolute values of these rate constants, but only by their ratio (in range from $10^{3.4}$ to 10^7 for k_{-3}).

Kinetic curves in Fig. 4 (left) for different values of k_3 and k_{-3} illustrate the fact that reaction rate practically does not depend on absolute values of rate constants of k_3 and k_{-3}. But this is not a property of this particular chemical mechanism. This phenomenon depends on initial reaction concentrations and can by discriminated by experimenter with an increase of $[E_{ox}]$ (from 10^{-9} to 10^{-8}, Fig. 4 (right)).

This research is focused on practical fittability of experimental data. Real experimental data have many issues, such as various experimental noise distributions, approximately known extinction coefficients or initial concentrations, not always precisely known reaction initial measurement time and others. These factors make subtle impact on data fitting, but of course, simultaneous effort to experiment design and data analysis can cope with these problems.

For example, in case of noisy data, where difference $x_{i,\exp} - x_{i,\text{calc}}$ are distributed normaly with zero mean value, we can estimate lowest fitting score value in case of biased curve. This can be done starting with expectation value definition [11]

$$\mathbf{E}\big[g(x)\big] \overset{\text{def}}{=} \int\limits_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x, \qquad (9)$$

where $g(x)$ is an arbitrary function of x, and $f(x)$ is a probability distribution function. Substitution to Eq. (9) our particular case with $g(x) = x^2$, $f(x) = \mathrm{e}^{-x^2/(2\sigma^2)}\sigma\sqrt{2\pi}$ and integration leads to

$$\sum_{i=1}^{n}(x_{i,\exp} - x_{i,\text{calc}})^2 \approx \int\limits_{-\infty}^{\infty} \frac{\mathrm{e}^{-x^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}x^2\,\mathrm{d}x = \sigma^2. \qquad (10)$$

This result suggest that ranking of solutions should not change dramatically, because the bias should only lift fitting score function by σ^2. Practically, it is not true. Adding to data noise with normal distribution ($\sigma^2 = 2 \cdot 10^{-6}$ and $\mu = 0$ – fairly reasonable parameters in experiment, example case #1 in Fig. 5) changes fitting results more than expected.
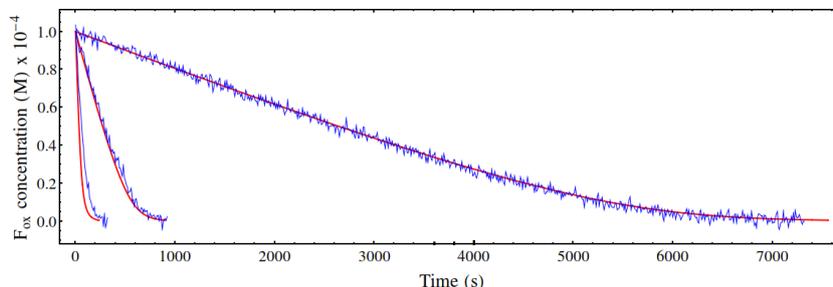
Fig. 5. Example of biased and not biased data for case #1.

Table 3. Fitting constants of three biased kinetic curves to system of differential equations (coarse $9 \times 9 \times 9 \times 9$ grid and bias from normal distribution $\sigma^2 = 2 \cdot 10^{-6}$ and $\mu = 0$).Initial concentrations $E_{ox} = 10^{-9}$, $M_{ox} = 10^{-(4 \div 6)}$, $F_{ox} = 10^{-4}$, $S = 10^{-3}$.

Case	Simulation constants				Worst Solution	Best solution
	$\lg(k_1)$	$\lg(k_2)$	$\lg(k_3)$	$\lg(k_{-3})$		
#1	6.3	7.3	6.3	5.3	fscore $= 2.95 \cdot 10^{-10}$	fscore $= 1.06 \cdot 10^{-11}$
					6.75 8.00 5.55 6.31	6.30 7.31 4.05 3.01
#2	5.3	7.3	6.3	5.3	fscore $= 3.35 \cdot 10^{-10}$	fscore $= 1.33 \cdot 10^{-11}$
					5.60 8.00 7.17 8.00	5.30 7.31 4.45 3.44
#3	5.3	6.3	6.3	5.3	fscore $= 4.19 \cdot 10^{-10}$	fscore $= 1.15 \cdot 10^{-11}$
					8.00 6.20 8.00 0.11	5.30 6.30 8.00 7.02
#4	5.3	6.3	5.3	5.3	fscore $= 2.04 \cdot 10^{-10}$	fscore $= 1.21 \cdot 10^{-11}$
					4.69 6.09 6.48 0.11	5.30 6.30 7.56 7.56
#5	5.3	6.3	5.3	4.3	fscore $= 4.32 \cdot 10^{-10}$	fscore $= 1.18 \cdot 10^{-11}$
					5.63 8.00 6.12 8.00	5.30 6.30 3.75 2.74
#6	5.0	8.0	7.0	7.0	fscore $= 4.17 \cdot 10^{-10}$	fscore $= 1.18 \cdot 10^{-11}$
					4.78 8.00 8.00 2.09	5.00 8.00 7.00 7.00

Fitting score lifted as expected. It could be checked simply by dividing fitting score by 3 (number of curves). Taking from Table 3 first row of biased curves fitting score $1.06 \cdot 10^{-11}$ and dividing by 3 and taking square root yields $1.87 \cdot 10^{-6}$, which is close to bias value of $\sigma = 2.0 \cdot 10^{-6}$. But as seen from same table, even by employing grid search it is not possible find absolute values of k_3 and k_{-3} (despite the fact, that k_1 and k_2 are fitted perfectly). This behaviour could be atributed to formed equilibrium in third reaction. As consequence of this equilibrium fitting score hypersurface exhibits narrow minima canyon, which is actually reached by the optimization procedure. In this case k_3 and k_{-3} are related to each other through equilibrium constant. Such phenomena are common in other contexts (for example in electrochemistry [4]).

5 Statistical analysis

Large set of solutions implies the necessity of statistical analysis. It could be splited in two parts – estimation of likelihood region and estimation of standard errors of fitting

parameters. The estimation of contour value of likelihood region for fitting score can be done with approximation from [1] and correction factor (in square brackets) from [11]:

$$c = \frac{s^2}{1 - I_{1-\alpha}^{-1}(P/2, (N-P)/2)} \left[\frac{\sum_{i=1}^m n_i/(n_i - 1)}{m} \right]. \tag{11}$$

Here N – total number of data points from all experimental curves, P – number of parameters, α – significance value, I^{-1} – inverse beta regularized function, s^2 – smallest fitting score value found on grid search, n_i – number of data points in certain experimental curve and m – number of experimental curves.

Inverse beta regularized function arise from Fisher–Snedecor cumulative distribution function, which is well known as $I_{xn/(xn+m)}(n, m)$, where I – beta regularized function, by solving equation $I_{xn/(xn+m)}(n, m) = 1 - \alpha$ for x [17]. Correction factor (in square brackets) arises from transformation of biased standard deviations into unbiased ones (with reasonable assumption that standard deviations are equal in all experimental curves). Formula can be used if standard deviations in all experiments are approximately same (as it should be true if noise level does not depend on signal level).

Estimation of errors is a byproduct of fitting. Summarized results from computation on grid with 0.25 logarithmic unit spacing and 20000 best optimized solutions are listed in Table 4.

Table 4. Final solutions of three biased curves fitting with estimated errors (with 95% confidence level).

Case	Data points	Critical fitting score contour value	Estimated mean and confidence interval (in logarithmic form)
#1	859	$1.0888 \cdot 10^{-11}$	$\lg(k_1) = 6.303\ [-0.010, +0.006]$ $\lg(k_2) = 7.303\ [-0.006, +0.016]$ $\lg(k_3) = 5.2\ [-1.4, +2.8],\quad \lg(k_{-3}) = 4.2\ [-1.5, +2.8]$ $\lg(K) = \lg(k_3/k_{-3}) = 1.02\ [-0.02, +0.05]$
#2	1057	$1.3518 \cdot 10^{-11}$	$\lg(k_1) = 5.300\ [-0.013, +0.003]$ $\lg(k_2) = 7.305\ [-0.004, +0.023]$ $\lg(k_3) = 5.3\ [-1.6, +2.7],\quad \lg(k_{-3}) = 4.3\ [-1.7, +2.7]$ $\lg(K) = \lg(k_3/k_{-3}) = 1.01\ [-0.02, +0.08]$
#3	8573	$1.1509 \cdot 10^{-11}$	$\lg(k_1) = 5.304\ [-0.002, +0.002]$ $\lg(k_2) = 6.302\ [-0.001, +0.002]$ $\lg(k_3) = 5.3\ [-1.6, +2.7],\quad \lg(k_{-3}) = 4.3\ [-1.6, +2.7]$ $\lg(K) = \lg(k_3/k_{-3}) = 0.98\ [-0.01, +0.01]$
#4	10514	$1.2144 \cdot 10^{-11}$	$\lg(k_1) = 5.297\ [-0.003, +0.002]$ $\lg(k_2) = 6.302\ [-0.002, +0.004]$ $\lg(k_3) = 4.7\ [-1.5, +3.3],\quad \lg(k_{-3}) = 4.7\ [-1.5, +3.3]$ $\lg(K) = \lg(k_3/k_{-3}) = -0.003\ [-0.003, +0.003]$
#5	8576	$1.1818 \cdot 10^{-11}$	$\lg(k_1) = 5.297\ [-0.003, +0.001]$ $\lg(k_2) = 6.301\ [-0.002, +0.004]$ $\lg(k_3) = 5.3\ [-2.0, +2.7],\quad \lg(k_{-3}) = 4.3\ [-2.0, +2.7]$ $\lg(K) = \lg(k_3/k_{-3}) = 1.011\ [-0.005, +0.011]$
#6	765	$1.1997 \cdot 10^{-11}$	$\lg(k_1) = 5.000\ [-0.001, +0.001],\quad \lg(k_2) = 8.000\ [0, 0]$ $\lg(k_3) = 5.5\ [-1.1, +2.5],\quad \lg(k_{-3}) = 5.5\ [-1.1, +2.5]$ $\lg(K) = \lg(k_3/k_{-3}) = 0.006\ [-0.006, +0.012]$

In all test cases, first two constants fit perfectly with small uncertainties. This is not valid for k_3 and k_{-3} – in Table 4 listed confidence intervals in logarithmic form shows uncertainties of several orders. Closer inspection of fitted k_3 and k_{-3} reveals perfect correlation between these constants (correlation coefficient $= 0.999972 \div 1$). This effect comes from chemical equilibrium formed in third reaction with particular initial concentrations and rate constants compensation in fitting score optimization. Perfect correlation between constants and chemical equilibrium implies dimensionality reduction of k_3 and k_{-3} by simply combining into equilibrium constant. Resulted equilibrium constant has much smaller uncertainties as a consequence of perfect correlation between k_3 and k_{-3}.

6 Conclusions

In this article, it was shown that the main reasons of possible failures to find desirable solutions are: the flat fitting score function surface; possible experimental design issues. We have shown that it is possible to deal with first problem by using exponential form of rate constants and starting optimization with evaluation of fitting score on constants grid and by selecting a set of possible candidates for further optimization, which leads to good fit or suggestion to modify experimental conditions. Also we showed how this method would cope with biased cases and how much information from certain experimental data can actually be extracted. Statistical analysis of fitting results identified confidence intervals of fitted parameters.

The approach suggested in our work is much more fruitful in experimental data fitting than the currently used strategies, which generate a single solution and offer no clue what is happening on entire fitting score function surface. Without such information it is impossible to draw any conclusions about quality and reliability of fit even when the fitting score looks excellent numerically.

The proposed method, where the final optimization step is performed on the set of grid searched initial points, is a general technique that can be applied to any case where optimization metric gives flat hypersurface, narrow canyons of minima, or other deviations from quadratic score function are expected.

References

1. D.M. Bates, D.G. Watts, Nonlinear regression: Iterative estimation and linear approximations, in *Nonlinear Regression Analysis and Its Applications*, John Wiley & Sons, 1988, pp. 32–66.

2. H. Bisswanger, Enzyme kinetics, in *Enzyme Kinetics: Principles and Methods*, Wiley-VCH, Weinheim, 2002, pp. 59–193.

3. W.E. Boyce, R.C. DiPrima, C.W. Haines, *Elementary Differential Equations and Boundary Value Problems*, Vol. 9, John Wiley & Sons, 1992.

4. A.D. Clegg, N.V. Rees, O.V. Klymenko, B.A. Coles, R.G. Compton, Experimental validation of Marcus theory for outer-sphere heterogeneous electron-transfer reactions: The oxidation of substituted 1,4-phenylenediamines. *ChemPhysChem*, **5**(8):1234–1240, 2004.

5. K.A. Connors, *Chemical Kinetics: The Study of Reaction Rates in Solution*. VCH Publishers, New York, Weinheim, Cambridge, 1990.

6. J. Kulys, Ž. Dapkūnas, The effectiveness of synergistic enzymatic reaction with limited mediator stability, *Nonlinear Anal. Model. Control*, **12**:495–501, 2007.

7. J. Kulys, L. Tetianec, I. Bratkovskaja, Pyrroloquinoline quinone-dependent carbohydrate dehydrogenase: Activity enhancement and the role of artificial electron acceptors, *Biotechnol. J.*, **5**(8):822–828, 2010.

8. P. Kuzmič, Dynafit – a software package for enzymology, *Methods Enzymol.*, **467**:247–280, 2009.

9. M. Molga, C. Smutnicki, Test functions for optimization needs, Technical Report, 2005.

10. O.V. Morozova, G.P. Shumakovich, S.V. Shleev, Ya.I. Yaropolov, Laccase-mediator systems and their applications: A review, *Appl. Biochem. Microbiol.*, **43**(5):523–535, 2007.

11. A. Papoulis, S.U. Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill Europe, 2002.

12. L. Petzold, Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations, *SIAM J. Sci. Stat. Comput.*, **4**(1):136–148, 1983.

13. M.J.D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comput. J.*, **7**(2):155–162, 1964.

14. G.E. Rovati, R. Shrager, S. Nicosia, P.J. Munson, Kinfit II: A nonlinear least-squares program for analysis of kinetic binding data, *Mol. Pharmacol.*, **50**(1):86–95, 1996.

15. D. Shoup, G. Lipari, A. Szabo, Diffusion-controlled bimolecular reaction rates. the effect of rotational diffusion and orientation constraints, *Biophys. J.*, **36**(3):697–714, 1981.

16. I.B. Svir, O.V. Klymenko, M.S. Platz, KINFITSIM – a software to fit kinetic data to a user selected mechanism, *Comput. Chem.*, **26**(4):379–386, 2002.

17. Wolfram Research, Inverse of the regularized incomplete beta function, `http://functions.wolfram.com/GammaBetaErf/InverseBetaRegularized/`, 2001–2008.