Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 108184, 13 pages
http://dx.doi.org/10.1155/2014/108184

Research Article

Numerical Reduced Variable Optimization Methods via Implicit Functional Dependence with Applications

Christopher Gunaseelan Jesudason

Department of Chemistry and Center for Theoretical and Computational Physics, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia

Correspondence should be addressed to Christopher Gunaseelan Jesudason; jesu@um.edu.my

Received 14 March 2014; Accepted 24 May 2014; Published 27 August 2014

Academic Editor: Mariano Torrisi

A systematic theoretical basis is developed that optimizes an arbitrary number of variables for (i) modeling data and (ii) the determination of stationary points of a function of several variables by the optimization of an auxiliary function of a single variable deemed the most significant on physical, experimental or mathematical grounds from which all the other optimized variables may be derived. Algorithms that focus on a reduced variable set avoid problems associated with multiple minima and maxima that arise because of the large numbers of parameters. For (i), both approximate *and* exact methods are presented, where the single controlling variable k of all the other variables $\mathbf{P}(k)$ passes through the local stationary point of the least squares metric. For (ii), an exact theory is developed whereby the solution of the optimized function of an independent variation of all parameters coincides with that due to single parameter optimization of an auxiliary function. The implicit function theorem has to be further qualified to arrive at this result. A nontrivial real world application of the above implicit methodology to rate constant and final concentration parameter determination is made to illustrate its utility. This work is more general than the reduction schemes for conditional linear parameters since it covers the nonconditional case as well and has potentially wide applicability.

1. Introduction

The following theory is a systematic development of all functions covering properties of constrained and unconstrained functions that are continuous and differentiable to various specified degrees [1, 2] and the proof of the existence of implicit functions [3] for the form of these functions to be optimized. The implicit function theorem is applied in a manner that requires further qualification because the optimization problem is of an unconstrained kind without any redundant variables. Methods (i)a,b (described in Sections 2 and 3, resp.) refer to modeling of data [4, Chapter 15, pages 773–806] where the form of the function $Q_{\mathrm{MD}}(\mathbf{P}, k)$ with independently varying variables $(\mathbf{P}, k)$ is

$$Q_{\mathrm{MD}}(\mathbf{P}, k) = \sum_{i=1}^{N'} \left(y_i - f(\mathbf{P}, t_i, k)\right)^2, \tag{1}$$

where y_i and t_i are datapoints and f is a known function, and optimizations of Q_{MD} may be termed a least squares (LS) fit over parameters $(\mathbf{P}, k)$ which are independently optimized for N' datasets. Method (ii) focuses on optimizing a general $Q_E(\mathbf{P}, k)$ function, not necessarily LS in form. There are many standard and hybrid methods to deal with such optimization [4, Chapter 10], such as golden section searches in 1D, simplex methods over multidimensions [4, pages 499–525], steepest descent and conjugate methods [5], and variable metric methods in multidimensions [4, pages 521–525]. Hybrid methods include multidimensional (DFP) secant methods [6], BFGS secant optimization [7], and RFO rational function optimization [8], which is a Newton-Raphson technique utilizing a rational function rather than a quadratic model for the function close to the solution point. Global deterministic optimization schemes combine several of the above approaches [9, Section 6.7.6]. Other physical methods, perhaps less easy to justify analytically, include probabilistic "basin-hopping" algorithms [9, Section 6.7.4], simulated annealing methods [10], and genetic algorithms [9, page 346]. An analytical justification on the other hand is

attempted here for these deterministic methods, but in real-world applications some of the assumptions (e.g., $\mathscr{C}^2$ continuity, compactness of spaces) may not always be obtained. For what follows, the distance metrics used are all Euclidean, represented by $|\cdot|$ or $\|\cdot\|$, where $\det\|\cdot\|$ represents the determinant of the matrix $\|\cdot\|$. Reduction of the number of variables to be optimized is possible in the standard matrix regression model only if conditional linear parameters $\boldsymbol{\beta}$ exist [11], where these $\boldsymbol{\beta}$ variables do not appear in the final $S(\boldsymbol{\theta})$ expression of the least squares function (2) to be optimized, whereas the $\boldsymbol{\phi}$ nonconditional linear parameters do and are a subset of the $\boldsymbol{\theta}$ variables; for the existence of each conditional linear parameter, there is a unit reduction in the number of independent parameters to be optimized. These reductions in variable number occur for any "expectation function" $f(\mathbf{x},\boldsymbol{\theta})$ which is the model or law for which a fitting is required, where there are N different datapoints $\mathbf{x}_i$, $i = 1, 2, \ldots, N$ that must be used to determine the p parameter variables $\boldsymbol{\theta}$ [11, page 32, Chapter 2]. A conditionally linear parameter θ_i exists if and only if the derivative of the expectation function $f(\mathbf{x},\boldsymbol{\theta})$ with respect to θ_i is independent of θ_i. Clearly such a condition may severely limit the number of parameters that can be neglected for the expectation function variables when the prescribed matrix regressional techniques are employed [11, Section 3.5.5, page 85] where the residual sum of squares is minimized:

$$S(\boldsymbol{\theta}) = \left\|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\right\|^2. \tag{2}$$

The N-vectors $\boldsymbol{\eta}(\boldsymbol{\theta})$ in P-dimensional space define the expectation surface. If the $\boldsymbol{\theta}$ variables are partitioned into the conditional linear parameters $\boldsymbol{\beta}$ and the other nonlinear parameters $\boldsymbol{\phi}$, then the response can be written $\boldsymbol{\eta}(\boldsymbol{\beta},\boldsymbol{\phi}) = \mathbf{A}(\boldsymbol{\phi})\boldsymbol{\beta}$. Golub and Pereyra [12] used a standard Gauss-Newton algorithm to minimise $S_2(\boldsymbol{\phi}) = \|\mathbf{y} - \mathbf{A}(\boldsymbol{\phi})\widehat{\boldsymbol{\beta}}(\boldsymbol{\phi})\|^2$ that depended only on the nonlinear parameters $\boldsymbol{\phi}$, where $\widehat{\boldsymbol{\beta}}(\boldsymbol{\phi}) = \mathbf{A}^+(\boldsymbol{\phi})\mathbf{y}$ with $\mathbf{A}^+$ being a defined pseudoinverse of $\mathbf{A}$ [11, Section 3.5.5, page 85], where $\mathbf{A}^+$ and $\mathbf{A}$ are matrices. The variables must be separable as discussed above and the number of variable reduction is only equal to the number of conditional linear parameters that exists for the problem. In applications, the preferred algorithm that exploits this valuable variable reduction is called variable projection. There are many applications in time resolved spectroscopy that is heavily dependent on this technique and many references to the method are given in the review by van Stokkum et al. [13]. Recently this method of variable projection has been extended in a restricted sense [14] in the field of inverse problems, which is not related to our method of either modeling or optimization, nor is the methodology related to the implicit function properties. In short, much of the reported methods developed are *ad hoc*, meaning that they are constructed to face the specific problems at hand with no claim to overall generality and this work too is *ad hoc* in the sense of suggesting variable reduction with specific classes of noninverse problems as indicated where the work develops a method of reducing the variable number to unity for all variables in the expectation function space *irrespective of whether they are conditional or not* by approximating their values by a method of averages (for method (i)a) without any form of linear regression being used in determining their approximations during the minimization iterations and without necessarily using the standard matrix theory that is valid for a very limited class of functions. Methods (i)b and (ii) are on the other hand exact treatments. No "elimination" of conditional linear parameters is involved in this nonlinear regression method. Nor is any projection in the mathematical sense involved. These general methods could have useful applications in deterministic systems comprising many parameters that are all linked to one variable: the primary one (denoted k here) that is considered on physical grounds to be the most significant. A generalization of this method would be to select a smaller set of variables than the full parameter list instead of just one variable as illustrated here. Another tool that could be used in conjunction with the reduced variable method is to employ the various search algorithms that has been actively developed to the reduction scheme developed here [15–19]. As far as we are aware, these optimization methods for multivariable problems all seem to mainly focus on various stochastic or deterministic methods using discrete algorithms in some type of search sequence of the domain space of the multivariable domain space $\{\mathbf{P}, k, t\}$ as detailed below in more recent publications. In other less ambitious works, the problem is narrowed to the specific nature of the system where the object function is specified and which is amenable to precise treatment, for example [20, 21], with a well-defined domain space, without variable reduction. In another context, quite different from the current development, variable reduction has been applied to DEA problems [22]. Other examples of multiparameter complex systems include those for multiple-step elementary reactions each with its own rate constant that gives rise to photochemical spectra signals that must be resolved unambiguously [23], but these belong to the class of functions with conditional linear parameters. The work here, on the other hand, predominantly focuses on accurately determining the range of the function to be optimized by reducing the space of the domain. Hence, this method can be successfully combined with the usual domain searching techniques mentioned above to effectively locate stationary points by a two-pronged approach. All these complex and coupled processes in physical theories are related by postulated laws $Y_{\text{law}}(\mathbf{P}, k, t)$ that feature parameters $(\mathbf{P}, k)$. Other examples include quantum chemical calculations with many topological and orientation variables that need to be optimized with respect to the energy, but in relation to one or a few variables, such as the molecular trajectory parameter during a chemical reaction where this variable is of primary significance in deciding on the "reasonableness" of the analysis [9, Section 6.2.3, page 294]. Methods (i)a and (i)b below refer to LS data-fitting algorithms. Method (i)a is an approximate method where it is proved under certain conditions; it could be a more accurate determination of parameters compared to a standard LS fit using (1). Method (i)b develops a technique where the optimum value for Q_{MD} with domain values $(\mathbf{P}, k)$ coincides with that of the standard LS method where the $(\mathbf{P}, k)$ variables are varied independently. Also discussed are the relative accuracy of both methods (i)a in Section 2.2

and (i)b (endnote at end of Section 3). Method (ii) develops a single parameter optimization where the conditions of an arbitrary $Q_{\mathrm{OPT}}(\mathbf{P},k)$ function are met simultaneously; namely,

$$\frac{\partial Q_{\mathrm{OPT}}(k)}{\partial k}=0\longrightarrow\left\{\frac{\partial Q_{\mathrm{OPT}}(\mathbf{P},k)}{\partial \mathbf{P}}=0,\frac{\partial Q_{\mathrm{OPT}}(\mathbf{P},k)}{\partial k}=0\right\}. \tag{3}$$

We note that methods (i)a, (i)b, and (ii) are not related to the Adomian decomposition method and its variants that expand polynomial coefficients [24] for solutions to differential equations not connected to estimation theory; indeed here there are no boundary values that determine the solution of the differential equations.

2. Method (i)a Theory

This approximate method utilizes the average of the N_c unique solutions for each value of k defined above, where the form of the fitting function—a "law" of nature for instance—is specified. Deterministic laws of nature are conveniently written in the form

$$Y_{\mathrm{law}}=Y_{\mathrm{law}}(\mathbf{P},k,t), \tag{4}$$

linking the variable Y_{law} to t. The components of $\mathbf{P}$, P_i ($i=1,2,\ldots,N_p$) and k are parameters. Verification of a law of form (4) relies on an experimental dataset $\{(Y_{\mathrm{exp}}(t_i),t_i),\ i=1,2,\ldots,N\}$. The t variable could be a vector of variable components of experimentally measured values or a single parameter as in the kinetics examples below where t_i denotes values of time t in the domain space. The vector form will be denoted by $\mathbf{x}$. Variables ($\mathbf{x}$) are defined as members of the "domain space" of the measurable system and similarly Y_{law} is the defined range or "response" space of the physical measurement. Confirmation or verification of the law is based on (a) deriving experimentally meaningful values for the parameters $(\mathbf{P},k)$ and (b) showing a good enough degree of fit between the experimental set $Y_{\mathrm{exp}}(t_i)$ and $Y_{\mathrm{law}}(t_i)$. In real world applications, to chemical kinetics, for instance, several methods [25–28] and so forth have been devised to determine the optimal $(\mathbf{P},k)$ parameters, but most if not all these methods consider the aforementioned parameters as autonomous and independent (e.g., [26]). A similar scenario broadly holds for current state-of-the-art applications of structural elucidation via energy functions [9, Chapsters 4, 6]. To preserve the viewpoint of the interrelationship between these parameters and the experimental data, we devise schemes that relate $\mathbf{P}$ to k for all P_i via the set $\{Y_{\mathrm{exp}}(t_i),t_i\}$ and optimize the fit over k-space only. That is there is induced a $P_i(k)$ dependency on k via the experimental set $\{Y_{\mathrm{exp}}(t_i),t_i\}$. The conditions that allow for this will also be stated for the different methods.

2.1. Details of Method (i)a. Let N_p be the number of components of the $\mathbf{P}$ parameter, N' the number of dataset pairs $(Y_{\mathrm{exp}}(t_i),t_i)$, and N_s the number of singularities where the use of a particular dataset (Y_{exp},t) leads to a singularity in the determination of $\overline{P}_i(k)$ as defined below and which must be excluded from being used in the determination of $\overline{P}_i(k)$. Then $(N_p+1)\le(N-N_s)$ for the unique determination of $\{\mathbf{P},k\}$. Let N_c be the total number of different datasets that can be chosen which does not lead to singularities. If the singularities are not choice dependent, that is, a particular dataset pair leads to singularities for all possible choices, then we have the following definition for N_c where ${}^{N-N_s}C_{N_p}=N_c$ is the total number of combinations of the data-sets $\{Y_{\mathrm{exp}}(t_i),t_i\}$ taken N_p at a time that does not lead to singularities in P_i. In general, N_c is determined by the nature of the datasets and the way in which the proposed equations are to be solved. Write Y_{law} in the form

$$Y_{\mathrm{law}}(t,k)=f(\mathbf{P},t,k), \tag{5}$$

and for a particular dataset $\{Y_{\mathrm{exp}}(t_i),t_i\}$, write $f(i)\equiv f(\mathbf{P},t_i,k)$. Define the vector function $\mathbf{f_g}$ with components $f_g(i)\equiv \mathrm{Y}_{\mathrm{exp}}(t_i)-f(i)=f_g(i)(\mathbf{P},k)$. Assume $\mathbf{f_g}\in\mathscr{C}^1$ defined on an open set K_0 that contains k_0.

Lemma 1. *For any k_0 such that* $\det\|\partial f_g(i)(\mathbf{P},k_0)/\partial P_j\|\neq 0$, *$\exists$ the unique function $\mathbf{P}(k)\in\mathscr{C}^1$(with components $P_i(k)\cdots P_{N_c}(k)$) defined on K_0 where $\mathbf{P}(k_0)=\mathbf{P_0}$, and where $\mathbf{f_g}(\mathbf{P}(k),k)=0$ for every $k\in K_0$.*

Proof. The above follows from the implicit function theorem (IFT) [3, Theorem 13.7, page 374] where $k\in K_0$ is the independent variable for the existence of the $\mathbf{P}(k)$ function. □

We seek the solutions for $\mathbf{P}(k)$ subject to the above conditions for our defined functions. Map $f\rightarrow Y_{\mathrm{th}}(\overline{\mathbf{P}},t,k)$ as follows:

$$Y_{\mathrm{th}}(t,k)=f\left(\overline{\mathbf{P}},t,k\right), \tag{6}$$

where the term $\overline{\mathbf{P}}$ and its components are defined below and where k is a varying parameter. For any of the $(i_1,i_2,\ldots,i_{N_p})$ combinations denoted by a combination variable $\alpha\equiv(i_1,i_2,\ldots,i_{N_p})$ where $i_j\equiv(Y_{\mathrm{exp}}(t_{i_j}),t_{i_j})$ is a particular dataset pair, it is in principle possible to solve for the components of $\overline{\mathbf{P}}$ in terms of k through the following simultaneous equations:

$$\begin{aligned}
Y_{\mathrm{exp}}\left(t_{i_1}\right)&=f\left(\mathbf{P},t_{i_1},k\right),\\
Y_{\mathrm{exp}}\left(t_{i_2}\right)&=f\left(\mathbf{P},t_{i_2},k\right),\\
&\vdots\\
Y_{\mathrm{exp}}\left(t_{i_{N_p}}\right)&=f\left(\mathbf{P},t_{i_{N_p}},k\right),
\end{aligned} \tag{7}$$

from Lemma 1. And each α choice yields a unique solution $P_i(k,\alpha)$ ($i=1,2,\ldots,N_p$), where $P_i(k,\alpha)\in\mathscr{C}^1$. Hence any function of $P_i(k,\alpha)$ involving addition and multiplication is also in $\mathscr{C}^1$. For each P_i, there will be N_c different solutions, $P_i(k,1),P_i(k,2),\ldots,P_i(k,N_c)$. We can define an arithmetic

mean (there are several possible mean definitions that can be utilized) for the components of $\overline{\mathbf{P}}$ as

$$\overline{P}_i(k) = \frac{1}{N_c}\sum_{j=1}^{N_c} P_i(k,j). \tag{8}$$

In choosing an appropriate functional form for $\overline{\mathbf{P}}$ (8) we assumed equal weightage for each of the dataset combinations; however, the choice is open, based on appropriate physical criteria. We verify below that the choice of $\overline{\mathbf{P}}(k)$ satisfies the constrained variation of the LS method so as to emphasize the connection between the level-surfaces of the unconstrained LS with the line function $\overline{\mathbf{P}}(k)$.

Each $P_i(k, j)$ is a function of k whose derivative is known either analytically or by numerical differentiation. To derive an optimized set, then for the LS method, define

$$Q(k) = \sum_{i=1'}^{N'}\left(Y_{\exp}(t_i) - Y_{\text{th}}(k,t_i)\right)^2. \tag{9}$$

Then for an optimized k, we have $Q'(k) = 0$. Defining

$$R(k) = c\cdot\sum_{i=1'}^{N'}\left(Y_{\exp}(t_i) - Y_{\text{th}}(k,t_i)\right)\cdot Y'_{\text{th}}(k,t_i), \tag{10}$$

the optimized solution of k corresponds to $R(k) = 0$ which has been reduced to a one-dimensional problem. The standard LS variation on the other hand states that the variables $P_T = \{\mathbf{P}, k\}$ in (5) are independently varied so that

$$Q_T(P_T) = \sum_{i=1}^{N'}\left(Y_{\exp}(t_i) - f(P_T,t_i)\right)^2, \tag{11}$$

with solutions for Q_T in terms of P_T whenever $\partial Q_T/\partial P_T = 0$. Of interest is the relationship between the single variable variation in (9) and the total variation in (11). Since $\overline{\mathbf{P}}$ is a function of k, then (11) is a constrained variation where

$$\delta Q(k) = \delta Q\left(\overline{\mathbf{P}}, k\right) = \left(\frac{\partial Q}{\partial \overline{\mathbf{P}}}\cdot\delta\overline{\mathbf{P}} + \frac{\partial Q}{\partial k}\right), \tag{12}$$

subjected to $g_i(\overline{\mathbf{P}}, k) = \overline{P}_i - h_i(k) = 0$ (i.e., $\overline{P}_i = h_i(k)$ for some function of k) and where $\overline{P}_i$ are the components of $\overline{\mathbf{P}}$. According to the Lagrange multiplier theory [3, Theorem 13.12, page 381] the function $f : \mathbb{R}^n \to \mathbb{R}$ has an optimal value at x_0 subject to the constraints $g : \mathbb{R}^n \to \mathbb{R}^m$ over the subset S where $\mathbf{g} = (g_1, g_2 \cdots g_m)$ vanishes; that is, $x_0 \in X_0$, where $X_0 = \{x : x \in S,\ \mathbf{g}(x) = 0\}$ when either of the following equivalent equations ((13), (14)) are satisfied:

$$D_r f(x_0) + \sum_{k=1}^{m}\lambda_k D_r g_k(x_0) = 0 \quad (r = 1,2,\ldots,n), \tag{13}$$

$$\nabla f(x_0) + \lambda_1\nabla g_1(x_0) + \cdots + \lambda_m\nabla g_m(x_0) = 0, \tag{14}$$

where $\det\|D_j g_i(x_0)\| \neq 0$ and the λ's are invariant real numbers. We refer to $\overline{P}_i$ as any variable that is a function of k constructed on physical or mathematical grounds, and not just to the special case defined in (8). Write

$$g_i = \overline{P}_i - h_i(k) = 0 \quad \left(i = 1,2,\ldots,N_p\right), \tag{15}$$

where $D_j g_i(x_0) = \delta_{ij}$ since $D_j = \partial/\partial\overline{P}_j$ and therefore $\det\|D_j g_i(x_0)\| \neq 0$. We abbreviate the functions $f(i) = f(\mathbf{P}, t_i, k)$ and $\overline{f}(i) = f(\overline{\mathbf{P}}, t_i, k)$. Define

$$f_Q(x) \equiv Q\left(\overline{\mathbf{P}}, k, t\right) = \sum_{i=1'}^{N'}\left(Y_{\exp}(t_i) - \overline{f}(i)\right)^2, \tag{16}$$

where $Y_{\exp}(t_i)$ are the experimental subspace variables as in (7) with $x \in X_0$ defined above. We next verify the relation between $Q(k)$ and Q_T.

Verification. The solution $Q'(k) = R(k) = 0$ of (10) is equivalent to the variation of $f_Q(x)$ defined in (16) subjected to constraints g_i of (15).

Proof. Define the Lagrangian to the problem as $\mathscr{L} = f_Q(x) + \sum_{i=1}^{N_p}\lambda_i g_i$. Then the equations that satisfy the stationary condition

$$\frac{\partial\mathscr{L}}{\partial\overline{P}_j} = 0, \quad j = 1,2,\ldots,N_p; \qquad \frac{\partial\mathscr{L}}{\partial k} = 0 \tag{17}$$

reduce to the (equivalent) simultaneous equations

$$\sum_{i=1}^{N'}\left(Y_{\exp}(t_i) - \overline{f}(i)\right)\frac{\partial\overline{f}(i)}{\partial\overline{P}_j} = \lambda'_j, \quad j = 1,2,\ldots,N_p, \tag{18}$$

$$\sum_{i=1}^{N'}\left(Y_{\exp}(t_i) - \overline{f}(i)\right)\frac{\partial\overline{f}(i)}{\partial k} + \sum_{j=1}^{N_p}\lambda'_j\frac{\partial p_j}{\partial k} = 0. \tag{19}$$

Substituting λ'_j in (18) to (19) leads to

$$\sum_{i=1}^{N'}\left(Y_{\exp}(t_i) - \overline{f}(i)\right)\frac{\partial\overline{f}(i)}{\partial k} + \sum_{j=1}^{N_p}\sum_{i=1}^{N'}\left(Y_{\exp}(t_i) - \overline{f}(i)\right)\cdot\frac{\partial\overline{f}(i)}{\partial\overline{P}_j}\cdot\frac{\partial p_j}{\partial k} = 0. \tag{20}$$

Since $d\overline{P}_i/dk = \partial\overline{p}_i/\partial k$, then (20) implies $dQ(\overline{\mathbf{P}}, k, t)/dk = 0$ for the Q functions in (11), (12), and (16). □

Of interest is the theoretical relationship of the $(\overline{\mathbf{P}}, k)$ variables of the Q functions described by (9), (12), (16) denoted Q_1 and those of the freely varying Q function of (1) denoted Q_2 with the variable set which can be written as

$$Q_1 = Q\left(\overline{\mathbf{P}}, t, k\right), \tag{21}$$

$$Q_2 = Q(\mathbf{P}, t, k), \tag{22}$$

which is given by the following theorem, where we abbreviate $\alpha_i = (Y_{\exp}(t_i) - f(\mathbf{P}, t_i, k))$ and $\overline{\alpha}_i = (Y_{\exp}(t_i) - f(\overline{\mathbf{P}}, t_i, k))$, where we note that the f functional form is unique and is of the same form for both these α variables.

Theorem 2. *The unconstrained LS solution to* $Q_2 = Q(\mathbf{P}, k, t)$ *for the independent variables* $\{\mathbf{P}, k\}$ *is also a solution for the constrained variation single variable* k', *where* $\mathbf{P} = \overline{\mathbf{P}}(k')$, $k = k'$. *Further, the two solutions coincide if and only if*

$$\sum_{i=1}^{N'} \left(Y_{\exp}(t_i) - f\left(\overline{\mathbf{P}}, t_i, k\right)\right) \frac{\partial f\left(\overline{\mathbf{P}}, t_i, k\right)}{\partial \overline{P}_j} = 0, \quad j = 1, 2, \ldots, N_p. \tag{23}$$

Proof. The Q_2 unconstrained solution is derived from the equations

$$\frac{\partial Q_2}{\partial P_j} = c \cdot \sum_{i=1}^{N'} \alpha_i \frac{\partial f(i)}{\partial P_j} = 0, \quad j = 1, 2, \ldots, N_p, \tag{24}$$

$$\frac{\partial Q_2}{\partial k} = c \cdot \sum_{i=1}^{N'} \alpha_i \frac{\partial f(i)}{\partial k} = 0, \tag{25}$$

with c being constants. If there is a $\overline{\mathbf{P}}(k)$ dependency, then we have

$$\frac{dQ_1(k)}{dk} = c \cdot \sum_{j=1}^{N_p} \left(\sum_{i=1}^{N'} \overline{\alpha}_i \frac{\partial \overline{f}(i)}{\partial P_j} \right) \frac{\partial \overline{P}_j}{\partial k} + c \cdot \sum_{i=1}^{N'} \overline{\alpha}_i \frac{\partial \overline{f}(i)}{\partial k}. \tag{26}$$

If the variable set $\{\mathbf{P}, k\}$ satisfies (24) and (25) in unconstrained variation, then the values when substituted into (26) satisfy the equation $dQ_1(k)/dk = 0$ since $f(i)$ and $\overline{f}(i)$ are the same functional form. This proves the first part of the theorem. The second part follows from the converse argument, where from (26), if $dQ_1(k)/dk = 0$, then setting one factor to zero in (27) leads to the implication of (28)

$$\sum_{i=1}^{N'} \overline{\alpha}_i \frac{\partial \overline{f}(i)}{\partial \overline{P}_j} = 0, \quad j = 1, 2, \ldots, N_p, \tag{27}$$

$$\Longrightarrow \sum_{i=1}^{N'} \overline{\alpha}_i \frac{\partial \overline{f}(i)}{\partial k} = 0, \tag{28}$$

which is the solution set $\{\overline{\mathbf{P}}(k'), k'\}$ which satisfies $dQ_1(k)/dk = 0$ and is satisfied by the conditions of both (27) and (28). Then (27) satisfies (24) and (28) satisfies (25). □

The theorem, verification, and lemma above do not indicate topologically under what conditions a coincidence of solutions for the constrained and unconstrained models exists. Figure 1 depicts the discussion below. From Theorem 2, if set A represents the solution $\{\mathbf{P}, k\}$ for the unconstrained LS method and set $B = \{\overline{\mathbf{P}}, k'\}$ for the constrained method, then $B \supseteq A$. Define k within the range $k_1 \le k \le k_2$. Then k is in a compact space, and since $\mathbf{P}(k) \in \mathscr{C}^1$, $\mathbf{P}(k)$ is uniformly continuous [2, Theorem 8, page 79]. Then admissible solutions to the above constraint problem with the inequality $B \supseteq A$ imply $Q(\mathbf{P}(k)) \ge Q_{\min}$, where $Q_{\min}$ is the unconstrained minimum. The unconstrained $Q = Q_T$ LS function to be minimized in (11) implies

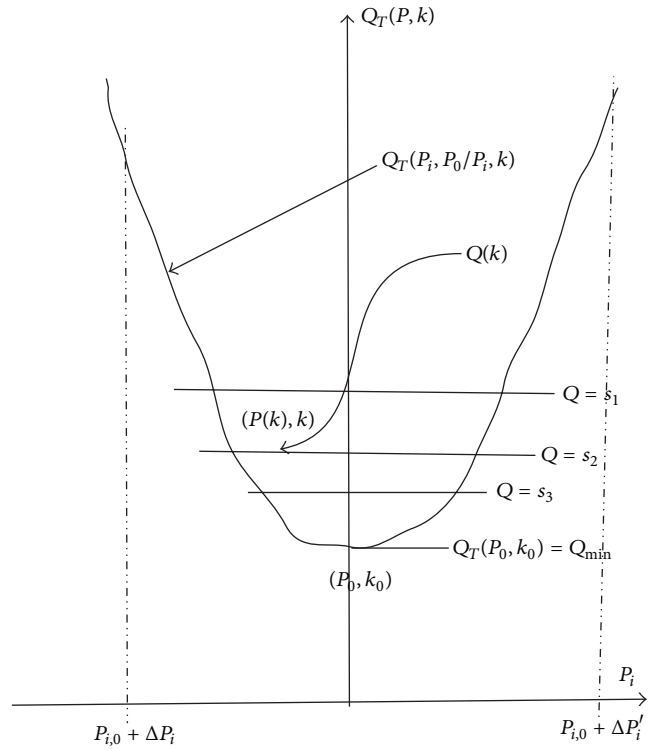

Figure 1: Depiction of how the k variation optimizing Q leads to a solution on a level surface of the Q_T function where $Q_T \le Q$.

$$\nabla Q = 0, \quad \left(\frac{\partial Q}{\partial k} = \frac{\partial Q}{\partial P_i} = 0, \text{ for } i = 1, 2, \ldots, N_p \right). \tag{29}$$

Defining the constrained function $Q_c(k) = \sum_{i=1}^{N'} (Y_{\exp}(i) - f(\mathbf{P}(k), t_i, k))^2$, then $Q_c(k) = Q \circ P_T$ where $P_T = (P_1(k), P_2(k), \ldots, P_{N_p}(k), k)^T$. Because $Q_c'(k) = (\partial Q/\partial \mathbf{P}, \partial Q/\partial k) \cdot (P_1'(k) \cdots P_{N_p}(k), k)^T$, solutions occur when (i) $\nabla Q = 0$ corresponding to the coincidence of the local minimum of the unconstrained Q for the best choice for the line with coordinates $(\mathbf{P}(k), k)$ as it passes through the local unconstrained minimum and (ii) $P_i'(k) = 0$, $(i = 1, 2, \ldots, N_p)$, $\partial Q/\partial k = 0$ where this solution is a special case of (iii) when the vector P_T' is $\perp$ to $\nabla Q \neq 0$; that is, P_T' is at a tangent to the surface $Q = S_2$ for some $S_2 \ge Q_{\min}$ where this situation is shown in Figure 1, where the vector is tangent at some point of the surface $Q_T = S_2$. Whilst the above characterizes the topology of a solution only, the existence of a solution for the line $(\mathbf{P}(k), k)$ which passes through the point of the unconstrained minimum of Q_T is proven below under certain conditions where a set of equations are constructed to allow for this significant application that specifies the conditions when the standard LS constrained variation solution implies the same solution as for the unconstrained variation. Also discussed is the case when it may be possible for unconstrained solution set U

to satisfy the inequality $Q_c(U) \geq Q_c$, where Q_c is a function designed to accommodate all solutions of (7), as given below in (30).

2.2. Discussion of LS Fit for a Function Q_c with a Possibility of a Smaller LS Deviation than for $\{\mathbf{P}, k\}$ Parameters Derived from a Free Variation of (11). The LS function metric such as (11) implied $Q(\mathbf{P}, k) \leq Q(\overline{\mathbf{P}}, k)$ at a stationary (minimum) point for variables $(\mathbf{P}, k)$. On the other hand, the sets of solutions of (7) denoted $\{\}_i$, N_c in number provides for each set exact solutions $\mathbf{P}(k)$ averaged to $\overline{\mathbf{P}}(k)$ using (8). If the $\{\mathbf{P}_i\}$, $i = 1, 2, \ldots, N_c$ solutions are in a δ-neighbourhood, then we examine the possibility that the composite function metric to be optimized over all the sets of equations $\{\}_i$, N_c in number defined here as

$$Q_c(\mathbf{P}, k) = \sum_{i \in \{\}_1, \{\}_2 \cdots \{\}_{N_c}} \left(Y_{\exp} - f(i)\right)^2, \tag{30}$$

could be such that

$$Q_c(\mathbf{P}_T, k) \geq Q_c\left(\overline{\mathbf{P}}, k\right), \tag{31}$$

where $(\mathbf{P}_T, k)$ is the unconstrained optimized value of (22). This implies that under these conditions, the Q_c of (30) is a better measure of fit. This will be proven to be the case under certain conditions below. For what follows, the $\mathbf{P}(k)_{\{\}_i}$ for equation set $\{\}_i$ obtains for all k values of the open set $\mathcal{S}$, $k \in \mathcal{S}$, from the IFT, including k_0 which minimises (9). Another possibility that will be discussed briefly later is where in (30), all $\{\mathbf{P}, k\}$ are free to vary. Here we consider the case of the N_c $\mathbf{P}$ values averaged to $\overline{\mathbf{P}}$ for some k. We recall the intermediate-value theorem (IVT) [3, Theorem 4.38, page 87] for real continuous functions f defined over a connected domain S which is a subset of some $\mathbb{R}^n$. We assume that the f functions immediately below obey the IVT. For each $\{\mathbf{P}_i\}$ solution of the $\{\}_i$ set for a specific $k = k_0$ we assume that the function $Q_{c,i}(\mathbf{P})$ is a strictly increasing function in the sense of definition (7) below, where

$$Q_{c,i}(\mathbf{P}) = \sum_{j \in \{\}_i} \left(Y_{\exp}(j) - f(j)\right)^2, \tag{32}$$

with $Q_{c,i}(\mathbf{P})(k_0) = 0$, in the following sense.

Definition 3. A real function f is (strictly) increasing in a connected domain $S \in \mathbb{R}^N$ about the origin at r_0 if relative to this origin, if $|r_2| > |r_1|$ (for the boundaries ∂ of ball $B(r_0, r_1)$ and $B(r_0, r_2)$) implies both $\max f(\partial B(r_0, r_2))(>) \geq \max f(\partial B(r_0, r_1))$ and $\min f(\partial B(r_0, r_2))(>) \geq \min f(\partial B(r_0, r_1))$.

Note. A similar definition is obtained for a (strictly) decreasing function with the $(<)$ $\leq$ inequalities. Since the ∂B boundaries are compact and f is continuous, the maximum and minimum values are attained for all ball boundaries. We assume f to be strictly increasing relative to r_0 for what follows below.

Lemma 4. *For any region bounded by $\partial B(r_0, r_1)$ and $\partial B(r_0, r_2)$ with coordinate* $\mathbf{r}$ *(radius r centered about coordinate r_0),*

$$\min f\left(\partial B\left(r_0, r_1\right)\right) < f(\mathbf{r}) < \max f\left(\partial B\left(r_0, r_2\right)\right). \tag{33}$$

Proof. Suppose in fact $f(\mathbf{r}) < \min f(\partial B(r_0, r_1))$; then

$$\min f\left(\partial B\left(r_0, r\right)\right) < \min f\left(\partial B\left(r_0, r_1\right)\right) \quad (r > r_1) \tag{34}$$

which is a contradiction to the definition and a similar proof is obtained for the upper bound. □

Note. Similar conditions apply for the nonstrict inequalities $\leq, \geq$.

The function that is optimized is

$$Q_c(\mathbf{P}) = \sum_{i=1}^{N_c} Q_{c,i}(\mathbf{P}). \tag{35}$$

Define $\mathbf{P}_i$ as the solution vector for the equation set $\{\}_i$. We illustrate the conditions where the solution $\mathbf{P}_T$ for a free variation for the Q metric given in (11) can fulfill the inequality where Q_c is as defined in (35)

$$Q_c(\mathbf{P}_T) \geq Q_c\left(\overline{\mathbf{P}}\right) \tag{36}$$

with $\overline{\mathbf{P}}$ given as in (8). A preliminary result is required. Define $\max \|\mathbf{P}_i - \mathbf{P}_j\| = \delta$, for all $i, j = 1, 2, \ldots, N_c$ and $\delta\mathbf{P}_i = \overline{\mathbf{P}} - \mathbf{P}_i$.

Lemma 5. $\|\delta\mathbf{P}_i\| \leq \delta$.

Proof.

$$\left\|\overline{\mathbf{P}} - \mathbf{P}_i\right\| = \frac{1}{N_c}\left\|\sum_q \mathbf{P}_q - N_c\mathbf{P}_i\right\| \leq \frac{1}{N_c}\sum_{j=1}^{N_c}\left\|\mathbf{P}_j - \mathbf{P}_i\right\| \leq \delta. \tag{37}$$

□

Lemma 6. *$Q_c(\mathbf{P}_T, k) > Q_c(\overline{\mathbf{P}}, k)$ for $\delta < \|\mathbf{P}_T - \mathbf{P}_i\| < \delta_T$ for some δ_T.*

Proof. Any point $\overline{\mathbf{P}}$ would be located within a spherical annulus centered at $\mathbf{P}_i$, with radii chosen so that by Lemma 4, we have the following results:

$$\epsilon_{\max,i} > Q_{c,i}\left(\overline{\mathbf{P}}\right) > \epsilon_{\min,i}, \tag{38}$$

where $f = Q_{c,i}$ in (32). Choose δ_i so that $\delta_i < \|\delta\mathbf{P}_i\| < \delta$. Define $\mathrm{Ann}(\delta, \delta_i, \mathbf{P}_i)$ as the space bounded by the boundary of the balls centered on $\mathbf{P}_i$ of radius δ and δ_i $(\delta > \delta_i)$. Then $\overline{\mathbf{P}} \in \mathrm{Ann}(\delta, \delta_i, \mathbf{P}_i)$ by Lemma 5. Since Q_c in (30) is not equivalent to $Q_T = Q$ in (11) where we write here the free variation vector solution as $\mathbf{P}_T$, then the above results lead to the following:

$$\sum_{i=1}^{N_c}\epsilon_{\min,i} < Q_c\left(\overline{\mathbf{P}}\right) = \sum_{i=1}^{N_c} Q_{c,i}\left(\overline{\mathbf{P}}\right) < \sum_{i=1}^{N_c}\epsilon_{\max,i}, \tag{39}$$

$$\epsilon_{\max,i} < Q_{c,i}\left(\mathbf{P}_T\right) < \epsilon_{T,i}, \tag{40}$$

where (40) follows from (33). Summing (40) leads to $Q_c(\mathbf{P}_T, k) > Q_c(\overline{\mathbf{P}}, k)$. □

Hence we have demonstrated that it may be more realistic or accurate to fit parameters based on a function that represents different coupling sets such as Q_c above rather than the standard LS method using (30) if $\mathbf{P}_T$ lies sufficiently far away from $\overline{\mathbf{P}}$. We note that if $\mathbf{P}_T$ is the solution of the free variation of the above Q_c in (30), then from the arguments presented after the proof of Theorem 2, it follows that

$$Q_c\left(\mathbf{P}_T\right) \le Q_c\left(\overline{\mathbf{P}}\right), \tag{41}$$

which implies that the independent variation of all parameters in LS optimization of the Q_c variation is the most accurate functional form to use assuming equal weighting of experimental measurements than the standard free variation of parameters using the Q function of (11).

3. Method (i)b Theory

Whilst it is advantageous in science data analysis to optimize a particular multiparameter function by focusing on a few key variables (our k variable of restricted dimensionality, which we have applied to a 1-dimensional optimization in the next section), it has been shown that this method yields a solution that is always of higher value for the same Q function than a full, independent parameter optimization, meaning that it is less accurate. The key issue, therefore, is whether for any Q function, including those of the Q_c variety, it is possible to construct a k parameter optimization such that the line of parameter variables $\mathbf{P}(k)$ passes through the minimum surface of the Q function. We develop a theory to construct such a function below. However, method (i)a may still be advantageous because of the greater simplicity of the equations to be solved, and the fact that $\mathscr{C}^1 f(i)$ functions were required, whereas here the $f(i)$ functions must be at least $\mathscr{C}^2$ continuous.

Theorem 7. *For the $Q_T(\mathbf{P},k)$ function defined in* (11)*, where each of the $f(i)$ functions is $\mathscr{C}^2$ on an open set $\mathbb{R}^{N_p+1}$ and where Q_T is convex, the solution at any point k of $\partial Q_T/\partial P_j = 0$, $(j = 1,2,\ldots,N_p)$ whenever $\det\|\partial \mathrm{Q_T}/\partial \mathrm{P_i}\partial \mathrm{P_j}\| \neq 0$ at $(dQ_T/dk)(k') = 0$ determines uniquely the line equation $\mathbf{P}(k)$ that passes the minimum of the function Q_T when $k = k'$.*

Proof. As before $f(i) = f(\mathbf{P}, t_i, k)$, so that

$$Q_T = Q(\mathbf{P},k) = \sum_{i=1}^{N'}\left(Y_{\exp(t_i)} - f(i)\right)^2. \tag{42}$$

Define $\partial f(i)/\partial P_j = f(i,j)$, $\alpha_k = Y_{\exp}(t_k) - f(k)$ and for an independent variation of the variables $(\mathbf{P},k)$ at the stationary point, we have

$$\frac{\partial Q_T}{\partial P_j} = h_j(\mathbf{P},k) = c\cdot\sum_{i=1}^{N'}\left(Y_{\exp}(t_i) - f(i)\right) f(i,j) = 0, \quad j = 1,2,\ldots,N_p, \tag{43}$$

$$\frac{\partial Q_T}{\partial k} = I(\mathbf{P},k) = c\cdot\sum_{i=1}^{N'}\alpha_i\frac{\partial f(i)}{\partial k} = 0. \tag{44}$$

The above results for the functions $h_j(\mathbf{P},k) = 0$ $(j = 1,2,\ldots,N_p)$ having a unique implicit function of k, denoted $\mathbf{P}(k)$ by the IFT [3, Theorem 13.7, page 374], require that $\det\|\partial h_i(\mathbf{P},k)/\partial P_j\| = \det\|\partial Q_T/\partial P_j\partial P_i\| \neq 0$ on an open set $S, k \in S$. More formally, the expansion of the preceding determinant in (45) verifies that a symmetric matrix is obtained for $\partial h_i(\mathbf{P},k)/\partial P_j$ due to the commutation of second-order partial derivatives of P_j

$$\frac{\partial h_i(\mathbf{P},k)}{\partial P_j} = c\cdot\sum_{l}^{N'}\left(\alpha_k\frac{\partial^2 f(l)}{\partial P_j\partial P_i} - \frac{\partial f(l)}{\partial P_j}\cdot\frac{\partial f(l)}{\partial P_i}\right). \tag{45}$$

Defining $Q_1(k)$ as a function of k only by expanding Q_T yields the total derivative with respect to k as $Q_1'(k)$, where

$$Q_1(k) = \sum_{i}^{N'}(\alpha_i)^2, \tag{46}$$

$$Q_1'(k) = c\cdot\sum_{j=1}^{N_p}\frac{dP_j}{dk}\cdot\sum_{i=1}^{N'}\alpha_i\frac{\partial f(i)}{\partial P_j} + c\cdot\sum_{i=1}^{N'}\alpha_i\frac{\partial f(i)}{\partial k}. \tag{47}$$

Then $h_i(\mathbf{P},k) = 0$ by construction (43) so that $\partial Q_T/\partial P_j = 0$ $(j = 1,2,\ldots,N_p)$ and (43) implies $\sum_{i=1}^{N'}\alpha_i(\partial f(i)/\partial P_j) = 0$ (for all j) and hence

$$\left(\frac{dP_j}{dk}\cdot\sum_{i=1}^{N'}\alpha_i\frac{\partial f(i)}{\partial P_j}\right) = 0. \tag{48}$$

Substituting (48) derived from (43) and (44) into (47) together with the condition $Q_1'(k) = 0$ implies that $c\cdot\sum_{i=1}^{N'}\alpha_i(\partial f(i)/\partial k) = 0$, which satisfies (44) for the free variation in k. Thus, $Q_1'(k) = 0 \Rightarrow \delta Q_T = 0$ for independent variation of $(\mathbf{P},k)$. So Q_T fulfills the criteria of a stationary point at say $k = k_0$, since $\nabla_{\mathbf{P},k}Q_T = 0$ ([2, Proposition 16, page 112]). Suppose that Q_T is convex, where $P_T = \{\mathbf{P}_0, k_0\}$ is a minimum point, $P_T \in D$, a convex subdomain of Q_T. Then at P_T, $\nabla_{\mathbf{P},k}Q_T = 0$, and P_T is also the unique global minimum over D according to [1, Theorem 3.2, page 46]. Thus P_T is unique, whether derived from a free variation of $(\mathbf{P},k)$ or via $P(k)$ dependent parameters with the Q_1 function. □

Note. As before, Q_T and Q_1 may be replaced with the summation of indexes as for Q_c in (30) to derive a physically more accurate fit.

4. Method (ii) Theory

Methods (i)a and (i)b which are mutual variants of each other are applications of the implicit method to modeling problems to provide a best fit to a function. Here, another variant of the implicit methodology for optimization of a target or cost function Q_E is presented. One can for instance consider $Q_E(\mathbf{P},k)$ to be an energy function with coordinates $\mathbf{R} = \{\mathbf{P},k\}$, where as before the components of $\mathbf{P}$ are P_j, $j = 1,2,\ldots,N_p$, $k \in \mathbb{R}$ is another coordinate so that $\mathbf{R} \in \mathbb{R}^{N_p+1}$.

For bounded systems, (such as the molecular coordinates), one can write

$$l_{\min,i} \le P_i \le l_{\max,i}, \quad i = 1, 2, \ldots, N_p; \qquad l_{\min,k} \le k \le l_{\max,k}. \tag{49}$$

Thus, $\mathbf{R} \in \mathscr{D} \subset \mathbb{R}^{N_p+1}$ is in a compact space $\mathscr{D}$. Define

$$\frac{\partial Q_E}{\partial P_j} = o_j(\mathbf{P}, k), \quad j = 1, 2, \ldots, N_p, \tag{50}$$

$$\frac{\partial Q_E}{\partial k} = \kappa(\mathbf{P}, k). \tag{51}$$

Then the equilibrium conditions become

$$o_j(\mathbf{P}, k) = 0, \qquad \kappa(\mathbf{P}, k) = 0. \tag{52}$$

Take (50) as the defining equations for $o_j(\mathbf{P}, k)$ which is specified by Q_E in (50) which casts it in a form compatible with the IFT where some further qualification is required for $(\mathbf{P}, k)$. Assume Q_E is $\mathscr{C}^2$ on $\mathscr{D}$, and $\det \|\partial o_i/\partial P_j\| \neq 0$, where $\partial o_i/\partial P_j \equiv \partial^2 Q_E/\partial P_j \partial P_i$. The matrix of the aforementioned determinant is symmetric, partaking of the properties due to this fact. Then by the IFT [3, Theorem 13.7, page 374], $\exists$ is a unique $\mathbf{P}(k)$ function where for some k_0, $o_j(\mathbf{P}(k_0), k_0) = 0$, $o_j \in \mathscr{C}^1$ on $\mathscr{T}_0$ with $(\mathscr{T}_0 \times \mathbb{R}) \subseteq \mathscr{D}$ and $k_0 \in \mathscr{T}_0$ such that $o_j(\mathbf{P}(k), k) = 0$ for all $k \in \mathscr{T}_0$. For k an isolated point $k = k_o$, from analysis, we find k to be still open. Write $Q_{E,1}(k) = Q_E(\mathbf{P}(k), k)$, and $Q'_{E,1}(k) = dQ_{E,1}/dk$, so that

$$\frac{dQ_{E,1}}{dk} = \sum_{j=1}^{N_p} \left(\frac{\partial Q_E}{\partial P_j} \cdot \frac{dP_j}{dk} + \frac{\partial Q_E}{\partial k} \right) = \sum_{j=1}^{N_p} \left(o_j \cdot \frac{dP_j}{dk} + \frac{\partial Q_E}{\partial k} \right). \tag{53}$$

Denote k_i as a solution to $Q'_{E,1}(k) = 0$, where $k_i \in \mathscr{T}_0$ in the indicated range above in (49).

Theorem 8. *The stationary points k_i $(i = 1, 2, \ldots, N_k)$ where $Q_{E,1}(k) = 0$ for $k = k_i$ exist for the range $\{k_i : k_{\min} \le k \le k_{\max}\}$ of coordinate k if and only if for each of these k_i, (i) $Q_E(\mathbf{P}, k) \in \mathscr{C}^2$ and (ii) $\det\|\partial Q_E(\mathbf{P}, k_i)/\partial P_j \partial P_i\| \neq 0$ (for all k_i, $i = 1, 2, \ldots, k_{\max}$). Each of these points $k_i \in \mathbb{R}^1$ space corresponds uniquely in a local sense in the open set $\mathscr{T}_0$ to some equilibrium (stationary) point of the target function $Q_E(\mathbf{P}, k)$ in $\mathbb{R}^{N_p+1}$ space.*

Proof. If $Q'_{E,1}(k) = 0$ where $k \in \mathscr{T}_0$, then it also follows from the IFT that $o_j = 0$, and therefore $\partial Q_E/\partial k = 0$ from (53), which satisfies (50) and (51) for the equilibrium point. The conditions (i) and (ii) of the theorem are a requirement of the IFT. Conversely, if $o_j = 0$, $(j = 1, 2, \ldots, N_p)$ and $\partial Q_E/\partial k = 0$ (a stationary or equilibrium point), then by (53) $Q'_{E,1}(k) = 0$. Hence, the coordinates $\{k_i\}$ for which $Q'_{E,1}(k_i) = 0$ refer to the condition, where $\delta Q_E(\mathbf{P}, k_i) = 0$, and uniqueness follows from the IFT reference to the local uniqueness of the $\mathbf{P}(k)$ function. □

Note. In a bounded system, one can choose any of the N_p components P_j of $\mathbf{P}$ as the k coordinate (denoted k_i), partly based on the convenience of solving the implicit equations to determine the k_i minima and thus determine by the uniqueness criterion the coordinates of the minima in $\mathbb{R}^{N_p+1}$ space (spanning the N_p independent variables and k).

For nondegenerate coordinate choice, meaning that for a particular k coordinate choice, there does not exist an equilibrium structure (meaning a set of coordinate values) where for any two structures A and B, $k_A = k_B$. For such structures, the total number of minima that exists within the bounded range in the k coordinate is equal to the total number of minima of the target function $Q_E(\mathbf{P}, k)$ within the bounded range. Hence, a method exists for the very challenging problem of locating and enumerating minima [9, Section 5.1, page 242 "How many stationary points are there?"]. From the uniqueness theorem of IFT, one could infer points in the k-axis where nonuniqueness is obtained; that is, whenever $\det \|\partial^2 Q_E/\partial P_i \partial P_j\| = 0$. In such cases, for particles with the same intermolecular potentials, permutation of the coordinates in conjunction with symmetry considerations could be of use in selecting the appropriate coordinate system to overcome these systems with degeneracies [9, Section 4.2.5, page 205, "Appearance and disappearance of symmetry elements"]. Other methods that might address this situation include scanning different one dimensional (1D) choices $k = P_i$ graphs or profiles, where if degeneracies exist for choice $k = P_j$, they may not exist for the choice $k = P_i$ in the graph for a specific point $(P_i \neq P_j)$. Thus, by scanning through all or selecting a number of the (1D) P_j profiles for $Q_{E,1}$, it would be possible to make an assignment of the location of a minimum in $\mathbb{R}^{N_p+1}$ space. One is reminded of the methods that spectroscopists use in assigning different energy bands based on selection rules to uniquely characterize, for instance, vibrational frequencies. A similar analogy is obtained for X-ray reflections, where the amplitude variation of the X-ray intensity in reciprocal space can be used to elucidate structure. The minima of the 1D, k coordinate scan must correspond to the minima in $\mathbb{R}^{N_p+1}$ space of the Q_E function given that all such minima in Q_E are locally strict and global within a small open set about the minima for by continuity, $Q_{E,1}(k) - Q_{E,1}(k_0) > 0$ for $|k - k_0| < \delta$ and for $|P(k) - P(k_0)| < \delta_2$, which violates the condition for a maximum.

5. Specic Algorithms and Pseudocode for Solution to Optimization Problem Utilizing Method (i)(a), Method (i)(b) and Method (ii)

We provide suggestions in pseudocode form for the above 3 proven methodologies. Real world applications of these methods are very involved undertakings that are separate research topics in their own right. Nevertheless, we provide a detailed and extended application of method (i)a suitable for a real world chemical kinetics problem where experimental data from the published literature are used for method (i)a in Section 5.3.1 and the results obtained compared to conventional techniques.

5.1. Pseudocode Algorithm for Method (i)a. Of the many variations possible, the following approach conforms to the theoretical development.

(1) For any physical law, for a total of N' datapoints, choose N_p datapoints (set α) and solve for parameters P_j, $j = 1, 2, \ldots, N_p$ according to (7) for each of the sets α, N_c in number for a known value of k. The solution set $\{\mathbf{P}(k)\}_\alpha$ may be derived analytically (as in the example below) or by appropriate linear approximations.

(2) Determine from the above set $\overline{\mathbf{P}}(k)$ either the geometric or statistical average (as used here) for the N_c solutions; that is, $\overline{\mathbf{P}}(k) = \sum_{\alpha=1}^{N_c} \mathbf{P}_\alpha(k)$.

(3) Determine $Q(k)$ (9) as

$$Q(k) = \sum_{i=1'}^{N'} \left(Y_{\text{exp}}(t_i) - Y_{\text{th}}(k, t_i)\right)^2. \tag{54}$$

(4) Solve the 1D equation at $k = k_0$ when $Q'(k) = 0$. The solution set is $\{\mathbf{P}(k_0), k_0\}$ for the optimization problem.

5.2. Pseudocode Algorithm for Method (i)b. This is an "exact" method relative to LS variation of all parameters. A suitable algorithm based on the theory could be as follows.

(1) Solve

$$h_j(\mathbf{P}, k) = c \cdot \sum_{i=1}^{N'} \left(Y_{\text{exp}}(t_i) - f(i)\right) f(i, j) = 0, \quad j = 1, 2, \ldots, N_p, \tag{55}$$

for a particular value of k. Since there are N_p equations for h_j, a solution $\mathbf{P}(k)$ exists. The solution may be exact or some linear approximation, depending on nature of the problem and convergence criteria.

(2) Form the function $Q_1(k) = \sum_i^{N'} (\alpha_i)^2$, (46), where $\alpha_i = Y_{\text{exp}}(t_i) - f(\mathbf{P}, t_i, k)$.

(3) Solve $Q_1'(k) = 0$ for some k_0; that is, $Q_1'(k_0) = 0$.

(4) The solution set to the problem is $\{\mathbf{P}(k_0), k_0\}$ for optimizing the LS function $Q_T = Q(\mathbf{P}, k) = \sum_{i=1}^{N'} (Y_{\text{exp}(t_i)} - f(\mathbf{P}, t_i, k))^2$.

5.3. Pseudocode Algorithm for Method (ii). Here $Q_E(\mathbf{P}, k)$ is a general function, not necessarily of form Q_T in (42). Then for N_p variables $\{\mathbf{P}\}$, define functions $o_j(\mathbf{P}, k) = \partial Q_E / \partial P_j$, $j = 1, 2, \ldots, N_p$ [22], $\kappa(\mathbf{P}, k) = \partial Q_E / \partial k$ as in (51).

(1) For a particular k, solve [22] for $\mathbf{P}$. $\mathbf{P}(k)$ exists since there are N_p equations. Approximate linearized solutions might also be attempted in the vicinity of the station point of k.

(2) Form the function $Q_{E,k}(k) = Q_E(\mathbf{P}, k)$.

(3) Solve for k_0, such that $Q_{E,k}'(k) = 0$.

(4) Solution to the optimization problem of $Q_E(\mathbf{P}, k)$ by varying independently all the domain variables is $\{\mathbf{P}(k_0), k_0\}$.

5.3.1. Application of Method (i)a Algorithm (Section 5.1) in Chemical Kinetics. The utility of one of the above triad of methods is illustrated in the determination of two parameters in chemical reaction rate studies, of 1st and 2nd order, respectively, using data from published literature, where method (i)a yields values close within experimental error to those quoted in the literature. The method can directly derive certain parameters like the final concentration terms (e.g., λ_∞ and Y_∞) if k, the rate constant, is the single optimizing variable in this approximation, which is not the case in most conventional methodologies. We assume here that the rate laws and rate constants are not slowly varying functions of the reactant or product concentrations, which have recently, from simulation, been shown to be generally not the case [29]. Under this standard assumption, the rate equations below are all obtained. The first-order reaction studied here is (i) the methanolysis of ionized phenyl salicylate with data derived from the literature [30, Table 7.1, page 381] and the second-order reaction analyzed is (ii) the reaction between plutonium(VI) and iron(II) according to the data in [31, Table II, page 1427] and [32, Tables 2–4, page 25].

5.3.2. First-Order Results. Reaction (i) above corresponds to

$$\text{PS}^- + \text{CH}_3\text{OH} \xrightarrow{k_a} \text{MS}^- + \text{PhOH} \tag{56}$$

where the rate law is pseudo first-order expressed as

$$\text{rate} = k_a[\text{PS}]^- = k_c\left[\text{CH}_3\text{OH}\right]\left[\text{PS}^-\right] \tag{57}$$

with the concentration of methanol held constant (80% v/v) and where the physical and thermodynamical conditions of the reaction appear in [30, Table 7.1, page 381]. The change in time t for any material property $\lambda(t)$, which in this case is the absorbance $A(t)$ (i.e., $A(t) \equiv \lambda(t)$) is given by

$$\lambda(t) = \lambda_\infty - (\lambda_\infty - \lambda_0)\exp(-k_a t) \tag{58}$$

for a first-order reaction where λ_0 refers to the measurable property value at time $t = 0$ and λ_∞ is the value at $t = \infty$ which is usually treated as a parameter to yield the best least squares fit even if its optimized value is less for monotonically increasing functions (for positive $d\lambda/dt$ at all t) than an experimentally determined $\lambda(t)$ at time t. In Table 7.1 of [30], for instance, $A(t = 2160\,\text{s}) = 0.897 > A_{\text{opt},\infty} = 0.882$ and this value of A_∞ is used to derive the best estimate of the rate constant as $16.5 \pm 0.1 \times 10^{-3}\,\text{sec}^{-1}$ in that work.

For this reaction, the P_i of (5) refers to λ_∞ so that $\mathbf{P} \equiv \lambda_\infty$ with $N_p = 1$ and $k \equiv k_a$. To determine the parameter λ_∞ as a function of k_a according to (9) based on the *entire* experimental $\{(\lambda_{\text{exp}}, t_i)\}$ dataset we invert (58) and write

$$\lambda_\infty(k) = \frac{1}{N'} \sum_{i=1'}^{N'} \frac{\left(\lambda_{\text{exp}}(t_i) - \lambda_0 \exp -kt_i\right)}{\left(1 - \exp -kt_i\right)}, \tag{59}$$

where the summation is for all terms with the i subscript of the experimental dataset that does not lead to zeros nor singularities, such as when $t_i = 0$. We define the nonoptimized, continuously deformable theoretical curve λ_{th}, where $\lambda_{\text{th}} \equiv Y_{\text{th}}(t,k)$ in (6) as

$$\lambda_{\text{th}}(t,k) = \lambda_{\infty}(k) - \left(\lambda_{\infty}(k) - \lambda_0\right)\exp\left(-k_a t\right). \quad (60)$$

With such a relationship of the λ_{∞} parameter P to k, we seek the least square minimum of $Q_1(k)$, where $Q_1(k) \equiv Q$ of (9) for this first-order rate constant k in the form

$$Q_1(k) = \sum_{i=1}^{N}\left(\lambda_{\exp}(t_i) - \lambda_{\text{th}}(t_i,k)\right)^2, \quad (61)$$

where the summation is over all the experimental $(\lambda_{\exp}(t_i), t_i)$ values. The solution of the rate constant k corresponds to the zero value of the function, which exists for both orders. The **P** parameters (λ_{∞} and Y_{∞}) are derived by back substitution into (59) and (65), respectively. The Newton-Raphson (NR) numerical procedure [4, page 456] was used to find the roots to P_k. For each dataset, there exists a value for λ_{∞} and so the error expressed as a standard deviation may be computed. The error tolerance for the NR procedure was set to 1.0×10^{-10}. We define the function deviation fd as the standard deviation of the experimental results with the best fit curve where $fd = \sqrt{(1/N)\{\sum_{i=1}^{N}(\lambda_{\exp}(t_i) - \lambda_{\text{th}}(t_i))^2\}}$. Our results are as follows: $k_a = 1.62 \pm .09 \times 10^{-2}\,\text{s}^{-1}$; $\lambda_{\infty} = 0.88665 \pm .006$; and $fd = 3.697 \times 10^{-3}$.

The experimental estimates are $k_a = 1.65 \pm .01 \times 10^{-2}\,\text{s}^{-1}$; $\lambda_{\infty} = 0.882 \pm 0.0$; and $fd = 8.563 \times 10^{-3}$.

The experimental method involves adjusting the $A_{\infty} \equiv \lambda_{\infty}$ to minimize the fd function and hence no estimate of the error in A_{∞} could be made. Method (i)a allows direct calculation of λ_{∞} and its error without the extraneous fittings required in the conventional methods. It is clear that our method has a lower fd value and is thus a better fit, and the parameter values can be considered to coincide with the experimental estimates within experimental error. Figure 2 shows the close fit between the curve due to our optimization procedure and experiment. The resulting R_k function (10) for the first-order reaction based on the published dataset is given in Figure 3. The very slight variation between the two curves could be due to experimental uncertainties as shown in Figure 1.

5.3.3. Second-Order Results. To further test our method, we also analyze the second-order reaction

$$\text{Pu (VI)} + 2\text{Fe (II)} \xrightarrow{k_b} \text{Pu (IV)} + 2\text{Fe (III)} \quad (62)$$

whose rate v is given by $v = k_0[\text{PuO}_2^{2+}][\text{Fe}^{2+}]$ where k_0 is relative to the constancy of other ions in solution such as H^+. The equations are very different in form to the first-order expressions and serves to confirm the viability of the current method.

For Espenson, the above stoichiometry is kinetically equivalent to the reaction scheme [32, equation (2-36)]

$$\text{PuO}_2^{2+} + \text{Fe}_{\text{aq}}^{2+} \xrightarrow{k_b} \text{PuO}_2^{+} + \text{Fe}_{\text{aq}}^{3+} \quad (63)$$

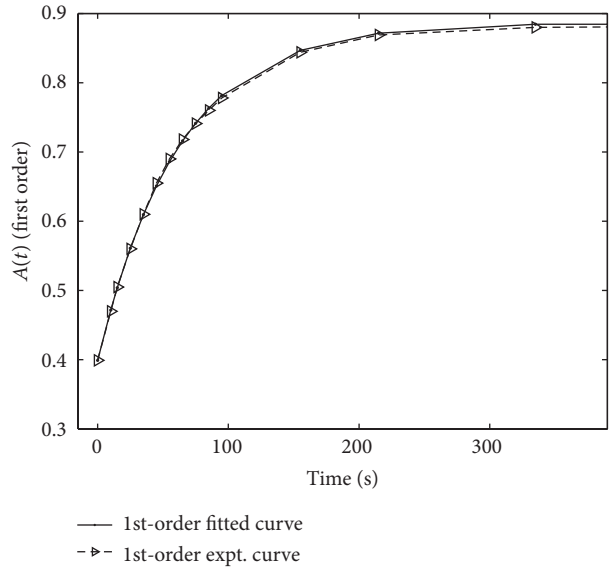

Figure 2: Plot of the experimental and curve with optimized parameters showing the very close fit between the two. The slight difference between the two can probably be attributed to experimental errors.

which also follows from the work of Newton and Baker [31, equations (8), (9), page 1429] whose data [31, Table II, page 1427] we use and analyze to verify the principles presented here. Espenson had also used the same data as we have to derive the rate constant and other parameters [32, pages 25-26] which are used to check the accuracy of our methodology. The overall absorbance in this case $Y(t)$ is given by [32, equation (2-35)]

$$Y(t) = \frac{Y_{\infty} + \{Y_0(1-\alpha) - Y_{\infty}\}\exp\left(-k\Delta_0 t\right)}{1 - \alpha\exp\left(-k\Delta_0 t\right)}, \quad (64)$$

where $\alpha = [\text{A}]_0/[\text{B}]_0$ is the ratio of initial concentrations where $[\text{B}]_0 > [\text{A}]_0$ and [B] = [Pu(VI)], [A] = [Fe(II)] and $[\text{B}]_0 = 4.47 \times 10^{-5}\,\text{M}$ and $[\text{A}]_0 = 3.82 \times 10^{-5}\,\text{M}$. A rearrangement of (64) leads to the equivalent expression [32, equation (2-34)]

$$\ln\left\{1 + \frac{\Delta_0(Y_0 - Y_{\infty})}{[\text{A}]_0(Y_t - Y_{\infty})}\right\} = \ln\frac{[\text{B}]_0}{[\text{A}]_0} + k\Delta_0 t. \quad (65)$$

According to Espenson, one cannot use this equivalent form [32, page 25] "because an experimental value of Y_{∞} was not reported" and he further asserts that if Y_{∞} is determined autonomously, then k, the rate constant, may be determined. Thus, central to all conventional methods is the autonomous and independent status of both k and Y_{∞}. We overcome this interpretation by defining Y_{∞} as a function of the total experimental spectrum of t_i values and k by inverting (64) to define $Y_{\infty}(k)$ as

$$Y_{\infty}(k) = \frac{1}{N'}\sum_{i=1'}^{N'}\frac{Y_{\exp}(t_i)\{\exp\left(k\Delta_0 t_i\right) - 1\} + Y_0(\alpha - 1)}{\left(\exp\left(k\Delta_0 t_i\right) - 1\right)}, \quad (66)$$

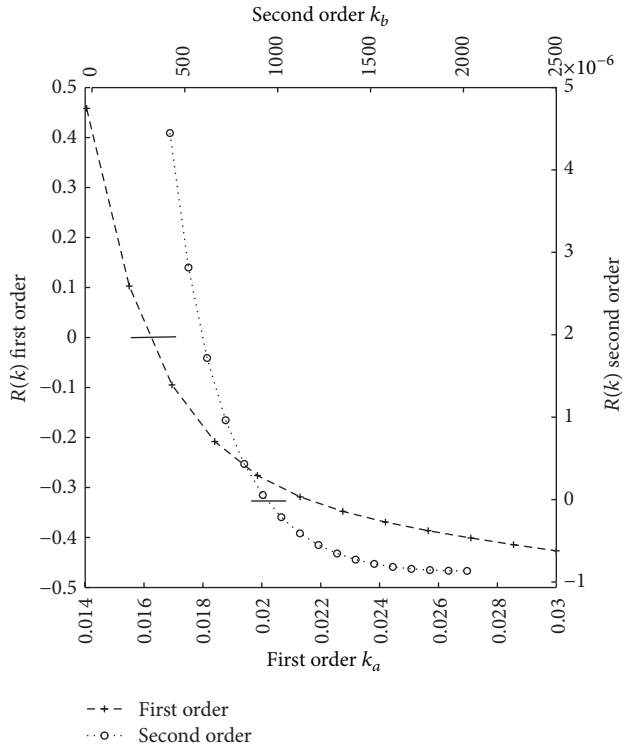

FIGURE 3: $R(k)$ functions (10) for reactions (i) and (ii) of order one and two in reaction rate. The horizontal bars correspond to the zero values of the $R(k)$ functions.

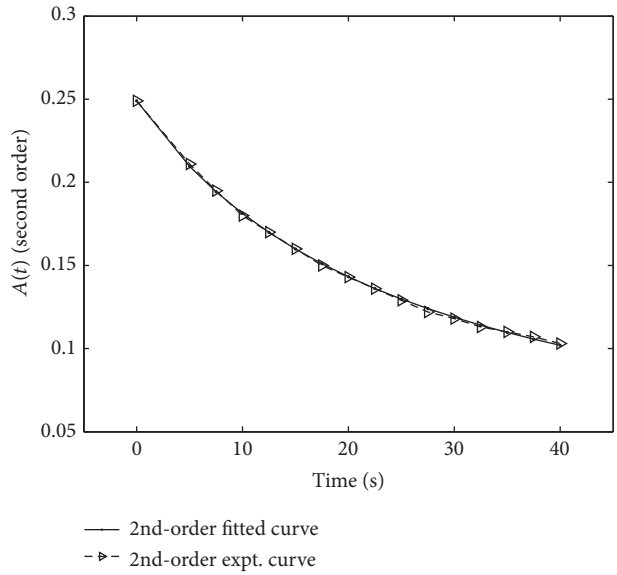

FIGURE 4: Graph of the experimental and calculated curve based on the current induced parameter-dependent optimization method.

where the summation is over all experimental values that does not lead to singularities such as at $t_i = 0$. In this case, the **P** parameter is given by $Y_\infty(k) = P_1(k)$, $k_b = k$ is the varying k parameter of (5). We likewise define a function Y_{th} of k that is also a function of t, but where the k parameter is interpreted as a "distortion" parameter in the following manner:

$$Y(t,k)_{\text{th}} = \frac{Y_\infty(k) + \{Y_0(1-\alpha) - Y_\infty(k)\}\exp(-k\Delta_0 t)}{1 - \alpha\exp(-k\Delta_0 t)}. \tag{67}$$

In order to extract the parameters k and Y_∞, we minimize the square function $Q_2(k)$ for this second-order rate constant with respect to k given as

$$Q_2(k) = \sum_{i=1}^{N}\left(Y_{\text{exp}}(t_i) - Y_{\text{th}}(t_i,k)\right)^2, \tag{68}$$

where the summation is over the N experimental t_i coordinates. Then the solution to the minimization problem is when the corresponding $R(k)$ function (10) is zero. The NR method was used to solve $R(k) = 0$ with the error tolerance of 1.0×10^{-10}. With the same notation as in the first-order case, the second order results are $k_b = 938.0 \pm 18\,(\text{Ms})^{-1}$; $Y_\infty = 0.0245 \pm 0.003$; and $fd = 9.606 \times 10^{-4}$.

The experimental estimates from the conventional methods are [32, page 25]: $k_b = 949.0 \pm 22\,(\text{Ms})^{-1}$; $Y_\infty = 0.025 \pm 0.003$.

Again the two results are in close agreement. The graph of the experimental curve and the one that is derived from our optimization method are given in Figure 4.

6. Conclusions

The triad of associated implicit function optimization covers both the topics of modeling of data *and* the optimization of arbitrary functions where experimental or theoretical considerations require that a single variable is tagged to a process variable that is iteratively relaxing to an equilibrium stationary point. Applying method (i)a to chemical kinetics allows for the direct determination of parameters that is not possible by application of the standard methodologies. The results presented here show that for linked variables, it is possible to derive all the parameters associated with a curve by considering only one independent variable which serves as the independent variable for other functions in the optimization process as illustrated by methods (i)a,b. Apart from possible reduced errors in the computations, it might also be a more accurate way of deriving parameters that are more influenced or conditioned (on physical grounds) by the value of one parameter (such as k here) than others; the current methods that give equal weight to all the variables might in some cases lead to results that would be considered "unphysical." In complex dynamical systems with multiprocesses, the physical considerations are such that for scientific purposes, it would be advantageous if optimization would be conducted on just one primary coordinate variable, such as in attempting to derive the most general stable conformer in a large molecule, where there are thousands of local minima present if all free coordinate variables are considered

[9, Section 6.7, page 330]. For such systems, method (ii) might be applicable. This generalized potential surface might be found suitable for reaction trajectory calculations [9, Chapter 4, page 192 on "Features of a landscape"] that require a single path variable, where the general optimized conformer would be relevant to the study of the potential surfaces and force fields present.

List of Variables

$(\mathbf{P}, t, k)$: the entire domain space of a function such as $f(\mathbf{P}, t, k)$ where k is the variable associated with a sequence of measurements, such as along the time coordinate. The vector $\mathbf{P}$ with components P_i is the normal parameter that must be optimized, and k is the specially chosen variable on experimental grounds that is optimized whilst constructing functions such that $\mathbf{P} = \mathbf{P}(k)$ (6)

$f(\mathbf{P}, t, k)$: refers to a function that is proposed to be a "law of nature" whose parameters $\{\mathbf{P}, k\}$ are to be optimized (7)

$Y_{\text{law}}(t, k)$: $Y_{\text{law}}(t, k) = f(\mathbf{P}, t, k)$ (5)

$Y_{\text{th}}(t, k)$: the theoretical law of nature if $\mathbf{P} = \overline{\mathbf{P}}(k)$ is determined. That is, $Y_{\text{th}}(t, k) = f(\overline{\mathbf{P}}, t, k)$ (6)

$Y_{\text{exp}}(t_i)$: an experimentally determined datapoint that ideally represents the range of $f(\mathbf{P}, t, k)$ if there was no error; that is, $Y_{\text{exp}}(t_i) = f(\mathbf{P}', t, k')$ for a perfect fit for all t_i, for fixed $(\mathbf{P}', k')$ (7)

$\overline{P}_i(k)$: an averaged value for $P_i(k)$ based on some specified algorithm (8)

$Q_X, Q(k)$: least squares (LS) function to optimize $Y_{\text{law}}(t, k)$ in the case of $Q(k)$ (9). In general, all Q_X functions are LS functions specified by X (e.g., (9), (11), (42), etc.)

Q_E: General cost or object function to be optimized, not necessarily in LS form (50)

$R(k)$: $= Q'(k)$ (10)

λ_i: Lagrange multipliers associated with the Q_X optimization (14)

$\mathscr{L}$: Lagrangian to the optimization problem (17)

k_q: chemical kinetics rate constant for reaction q (e.g., (56))

$\lambda(t), \lambda_\infty$: absorbance measurements for first-order chemical kinetics reactions at time t and infinity (e.g., (58) and (59))

$Y(t), Y_\infty$: absorbance measurements for second-order chemical kinetics reactions at time t and infinity (e.g., (64) and (66)).

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by University of Malaya Grant UMRG(RG077/09AFR) and Malaysian Government grant FRGS(FP084/2010A). This work was initiated and completed during a Sabbatical research visit (2012-2013) to the Atomistic Simulation Centre (ASC), School of Mathematics and Physics, Queen's University Belfast. I thank Ruth Lynden-Bell (Chemistry Department, Cambridge University) for facilitating this visit. Cordial discussions concerning real world applications with faculty at ASC are gratefully acknowledged. I thank my hosts, Jorge Kohanoff (ASC) and Christopher Hardacre (Chem. Dept., QUB) for congenial hospitality during this time.

References

[1] B. D. Craven, *Functions of Several Variables*, Chapman & Hall, London, UK, 1981.

[2] J. DePree and C. Swartz, *Introduction to Real Analysis*, John Wiley & Sons, New York, NY, USA, 1988.

[3] T. M. Apostol, *Mathematical Analysis*, Narosa Publishing House, New Delhi, India, 2nd edition, 2002.

[4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes—The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 3rd edition, 2007.

[5] J. A. Snyman, *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, Springer, New York, NY, USA, 2005.

[6] W. C. Davidon, "Variable metric method for minimization," *SIAM Journal on Optimization*, vol. 1, no. 1, pp. 1–17, 1991.

[7] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms," *Journal of the Institute of Mathematics and Its Applications*, vol. 6, pp. 76–90, 1970.

[8] A. Banerjee, N. Adams, J. Simons, and R. Shepard, "Search for stationary points on surfaces," *Journal of Physical Chemistry*, vol. 89, no. 1, pp. 52–57, 1985.

[9] D. A. Wales, *Energy Landscapes*, Cambridge Molecular Science, Cambridge University Press, Cambridge, UK, 2003, edited by R. Saykally, A. Zewail and D. King.

[10] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *The American Association for the Advancement of Science. Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[11] D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, NY, USA, 1988.

[12] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, pp. 413–432, 1973.

[13] I. H. M. van Stokkum, D. S. Larsen, and R. van Grondelle, "Global and target analysis of time-resolved spectra," *Biochimica et Biophysica Acta—Bioenergetics*, vol. 1657, no. 2-3, pp. 82–104, 2004.

[14] P. Shearer and A. C. Gilbert, "A generalization of variable elimination for separable inverse problems beyond least squares," *Inverse Problems*, vol. 29, no. 4, Article ID 045003, 27 pages, 2013.

[15] N. I. M. Gould, Y. Loh, and D. P. Robinson, "A filter method with unified step computation for nonlinear optimization," *SIAM Journal on Optimization*, vol. 24, no. 1, pp. 175–209, 2014.

[16] A. Milzarek and M. Ulbrich, "A semismooth Newton method with multidimensional filter globalization for l_1-optimization," *SIAM Journal on Optimization*, vol. 24, no. 1, pp. 298–333, 2014.

[17] H. M. Gomes, "Truss optimization with dynamic constraints using a particle swarm algorithm," *Expert Systems with Applications*, vol. 38, no. 1, pp. 957–968, 2011.

[18] M. Gendreau and J.-Y. Potvin, Eds., *Handbook of Metaheuristics*, vol. 146 of *International Series in Operations Research & Management Science*, Springer, New York, NY, USA, 2nd edition, 2010.

[19] R. Varadhan and P. D. Gilbert, "BB: an R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function," *Journal of Statistical Software*, vol. 32, no. 4, pp. 1–26, 2009.

[20] F. M. P. Raupp, L. M. G. Drummond, and B. F. Svaiter, "A quadratically convergent Newton method for vector optimization," *Optimization*, vol. 63, no. 5, pp. 661–677, 2014.

[21] M. D. Al-Khaleel, M. J. Gander, and A. E. Ruehli, "Optimization of transmission conditions in waveform relaxation techniques for RC circuits," *SIAM Journal on Numerical Analysis*, vol. 52, no. 2, pp. 1076–1101, 2014.

[22] A. Amirteimoori, D. K. Despotis, and S. Kordrostami, "Variables reduction in data envelopment analysis," *Optimization*, vol. 63, no. 5, pp. 735–745, 2014.

[23] S. Solar, W. Solar, and N. Getoff, "A pulse radiolysis-computer simulation method for resolving of complex kinetics and spectra," *Radiation Physics and Chemistry*, vol. 21, no. 1-2, pp. 129–138, 1983.

[24] A.-M. Wazwaz, "A note on using Adomian decomposition method for solving boundary value problems," *Foundations of Physics Letters*, vol. 13, no. 5, pp. 493–498, 2000.

[25] J. J. Houser, "Estimation of A_∞ in reaction-rate studies," *Journal of Chemical Education*, vol. 59, no. 9, pp. 776–777, 1982.

[26] P. Moore, "Analysis of kinetic data for a first-order reaction with unknown initial and final readings by the method of non-linear least squares," *Journal of the Chemical Society, Faraday Transactions I: Physical Chemistry in Condensed Phases*, vol. 68, pp. 1890–1893, 1972.

[27] W. E. Wentworth, "Rigorous least squares adjustment: application to some non-linear equations, I," *Journal of Chemical Education*, vol. 42, no. 2, pp. 96–103, 1965.

[28] W. E. Wentworth, "Rigorous least squares adjustment: application to some non-linear equations, II," *Journal of Chemical Education*, vol. 42, no. 3, pp. 162–167, 1965.

[29] C. G. Jesudason, "The form of the rate constant for elementary reactions at equilibrium from MD: framework and proposals for thermokinetics," *Journal of Mathematical Chemistry*, vol. 43, no. 3, pp. 976–1023, 2008.

[30] M. N. Khan, *Micellar Catalysis*, vol. 133 of *Surfactant Science Series*, Taylor & Francis, Boca Raton, Fla, USA, 2007, edited by A. T. Hubbard.

[31] T. W. Newton and F. B. Baker, "The kinetics of the reaction between plutonium(VI) and iron(II)," *Journal of Physical Chemistry*, vol. 67, no. 7, pp. 1425–1432, 1963.

[32] J. H. Espenson, *Chemical Kinetics and Reaction Mechanisms*, vol. 102, McGraw-Hill, Singapore, 2nd edition, 1995.