Trabajo Fin de Máster

MINLP in Chemical Reaction Networks

Presented by:
Mª Asunción Jiménez Cordero

Supervisors:
Dr. Rafael Blanquero Bravo, Universidad de Sevilla
Dr. Emilio Carrizosa Priego, Universidad de Sevilla

UNIVERSIDAD D SEVILLA

June 24, 2014

Contents

Chapter 1

Introduction

1.1 Problem Definition

A material balance for N chemical species taking place in L chemical reactions comprising a chemical reaction network can be written for certain classes of chemical reactors such as batch, fed-batch (semibatch) and continuous flow stirred tank reactor (CSTRs) ([19], [1], [10], [12], [18]). These balances are often written to reflect the pragmatic assumptions that the (homogeneous phase) reactor is operating isothermally, is well mixed, and that the overall density of the reaction mixture is not significantly changed by the occurrence of the chemical reactions within the reactor. The material balance expression for each species may be written in a very abstract form as (1.1):

$$\frac{dy_i}{dt} = f_i(y_1, \ldots, y_N, t, \theta), \quad i = 1, \ldots, N \tag{1.1}$$

where y_i is the molar concentration of specie E_i at time t, and θ is a multidimensional parameter. Equation (1.1) is a set of coupled ordinary differential equations (ODEs) that describe the dynamic behaviour of the reactive species due to chemical reactions, as represented by the N reaction fluxes f_i.

1.1.1 Stoichiometric Matrix and the Action Mass Law

In this work we focus on a particular form of (1.1) studying the so called action mass law. In this model, the flux terms f_i in (1.1) are directly linked to the stoichiometries of the L reactions taking place and the kinetic rate terms of these reactions. To be more precise, we consider the $N \times L$ stoichiometric

matrix $\mathbf{A}$ containing the stoichiometric coefficients for each of the N species in the L reactions:

$$\mathbf{A}_{N\times L} = \begin{pmatrix} a_1^1 & \cdots & a_1^L \\ \vdots & \ddots & \vdots \\ a_N^1 & \cdots & a_N^L \end{pmatrix}$$

Here, a_i^l is the stoichiometric coefficient of the chemical species E_i in the $l-$th reaction. By convention, $a_i^l < 0$ for a species that undergoes net consumption in a reaction (reactives), $a_i^l > 0$ for a species that undergoes net production (products), and $a_i^l = 0$ for a species that is either not involved in the $l-$th reaction or has no net change in it.

In the action mass law, reactives and products play very different roles. For this reason, it is convenient to express the stoichiometric matrix $\mathbf{A}$ as the difference of two matrices, $\mathbf{P}$, corresponding to the products, and $\mathbf{R}$, corresponding to the reactives,

$$\mathbf{A}_{N\times L} = \mathbf{P}_{N\times L} - \mathbf{R}_{N\times L} = \begin{pmatrix} \lambda_1^1 - \nu_1^1 & \cdots & \lambda_1^L - \nu_1^L \\ \vdots & \ddots & \vdots \\ \lambda_N^1 - \nu_N^1 & \cdots & \lambda_N^L - \nu_N^L \end{pmatrix},$$

where $a_i^l = \lambda_i^l - \nu_i^l$, $i = 1, \ldots, N$, $l = 1, \ldots, L$ and,

$$\mathbf{P}_{N\times L} = \begin{pmatrix} \lambda_1^1 & \cdots & \lambda_1^L \\ \vdots & \ddots & \vdots \\ \lambda_N^1 & \cdots & \lambda_N^L \end{pmatrix}; \qquad \mathbf{R}_{N\times L} = \begin{pmatrix} \nu_1^1 & \cdots & \nu_1^L \\ \vdots & \ddots & \vdots \\ \nu_N^1 & \cdots & \nu_N^L \end{pmatrix}$$

Example 1.1.1 *Let us consider a hypothetical chemical reaction network comprising four reactive species $E_1, \ldots, E_4$ ($N = 4$) involved in two elementary reactions ($L = 2$)*

$$\begin{aligned} E_1 + E_2 &\longrightarrow E_3 + E_4 \\ E_3 &\longrightarrow E_4 \end{aligned} \tag{1.2}$$

The stoichiometric matrix $\mathbf{A}$ *for this network is*

$$\mathbf{A}_{4\times 2} = \begin{pmatrix} -1 & 0 \\ -1 & 0 \\ 1 & -1 \\ 1 & 1 \end{pmatrix},$$

and the matrices $\mathbf{P}$ *and* $\mathbf{R}$ *corresponding to the products and reactives respectively are*

$$\mathbf{P}_{4\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}; \qquad \mathbf{R}_{4\times 2} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The component balances may now be conveniently expressed in terms of $\mathbf{A}$ and the L individual reaction rates in matrix-vector form:

$$\frac{d\mathbf{y}}{dt} = \mathbf{A}\delta(\mathbf{y}, t, \theta), \tag{1.3}$$

where δ is the $L \times 1$ dimensional vector of reaction rates, θ is a set of parameters, and $\mathbf{y}$ is the $N \times 1$ dimensional vector of species concentrations y_1,..., y_N at time t.

The L reactions involved in δ are, in general, nonlinear functions of the concentrations, $\mathbf{y}$, and linear functions of what is named the rate coefficients, $\mathbf{k}$. If elementary reactions are assumed, then the form of the L rate terms in δ is determined uniquely by the reactives in each of the L elementary equations.

The law of mass action kinetics states (see [19]) that the rate of an elementary reaction may be assumed to be directly proportional to the collision frequency of the reactives, and hence the product of the reactive concentrations; i.e. for a general reaction l with N species $y_1, \ldots, y_N$ and associated stoichiometric coefficients $\nu_1^l, \ldots \nu_N^l$ (reactives) and $\lambda_1^l, \ldots \lambda_N^l$ (products),

$$\sum_{i=1}^{N} \nu_i^l y_i \xrightarrow{k_l} \sum_{i=1}^{N} \lambda_i^l y_i,$$

the corresponding reaction rate δ is given by

$$\delta_l = k_l \prod_{i=1}^{N} y_i^{\nu_i^l}, \quad \text{for } l = 1, \ldots, L.$$

For the reaction network in (1.2), we can add the values of the rate coefficients in $\mathbf{k}$:

$$\begin{aligned} E_1 + E_2 &\xrightarrow{k_1} E_3 + E_4 \\ E_3 &\xrightarrow{k_2} E_4 \end{aligned} \tag{1.4}$$

The rate vector δ is in (1.5).

$$\delta = \begin{pmatrix} k_1 y_1 y_2 \\ k_2 y_3 \end{pmatrix} \tag{1.5}$$

Hence, the N ODEs, for known initial conditions, $y_i(0)$ of the concentrations $y_i(t)$, $i = 1, \ldots N$ can be used to describe the temporal evolution of the species concentrations:

$$\frac{d\mathbf{y}}{dt} = \mathbf{A}\delta(\mathbf{y}, t, \mathbf{k}) = \begin{pmatrix} -k_1 y_1 y_2 \\ -k_1 y_1 y_2 \\ k_1 y_1 y_2 - k_2 y_3 \\ k_1 y_1 y_2 + k_2 y_3 \end{pmatrix}$$

Note that any given row of $\mathbf{A}$ does not uniquely correspond to an elementary reaction, nor does any $\mathbf{A}$ correspond to a unique set of elementary reactions. For instance, the following reaction network has the same $\mathbf{A}$ as the network described by Equations (1.4):

$$E_1 + E_2 \xrightarrow{k_1} E_3 + E_4 \tag{1.6a}$$

$$E_1 + E_3 \xrightarrow{k_2} E_1 + E_4 \tag{1.6b}$$

1.2 Objective

Once we know which are the most important issues in our problem, let us see the goal. The objective of this problem is the following: given a set of empirical concentrations measured in some time instants, find the stoichiometric coefficients of the matrix $\mathbf{A}$ and the vector of rate coefficients $\mathbf{k}$ that best fits to the empirical concentrations. In other words, we want to find the best set of parameters, $\mathbf{A}, \mathbf{k}$, such that the concentrations, $\mathbf{y}$, obtained from (1.3) give the best fit with respect to the empirical concentrations.

1.3 Methodology

Let us make a summary of the most important steps in the problem resolution. In chapters 2-5 a more detailed explanation is given.

The first step is to build a function that fits the empirical concentrations in order to obtain also an approximation to the derivative of the empirical values. Let us recall that in this problem it is very important to know the concentrations but also the derivatives, since we have an ODE, (1.3), to be solved. For this purpose, different models are presented in Chapter 2.

Once the empirical data are smoothed to a given fitting function, the following step is to formulate the fitting problem, keeping in mind all the chemical constraints and giving a mathematical formula to express them. It is important to note that, as will be seen later, we have to solve a Mixed Integer Non Linear Problem (MINLP), where each component of the stoichiometric matrix $\mathbf{A}$ can only take integer values, and the components of the vector $\mathbf{k}$ are non negative real numbers.

A mixed integer nonlinear program (MINLP), [4], is a problem of the following form:

$$\begin{cases} \min\limits_{x,y} f_0(x,y) \\ s.t.\ f_j(x,y) \leq 0, \quad j = 1, \dots J \\ \qquad x \in \mathbb{Z}_+^n \\ \qquad y \in \mathbb{R}_+^m \end{cases} \tag{1.7}$$

where n is the number of integer-constrained variables, m is the number of continuous variables, J is the number of constraints, and finally,
$f_j : \mathbb{Z}_+^n \times \mathbb{R}_+^m \to \mathbb{R}, \quad j = 1, \ldots, J$ are arbitrary functions mapping.

In the next step, for a fixed stoichiometric matrix, $\mathbf{A}$, we begin a two-phase method in which the coefficient rate, $\mathbf{k}$ will try to be determined. These phases, called the differential method and the integral method, consider the problems (1.8) and (1.9), respectively, where e_{diff} and e_{int} are two different error measures detailed in Chapter 3 and $\mathfrak{K}$ is the set of constraints that the rate reaction vector $\mathbf{k}$ must obey:

$$\left\{ \begin{array}{ll} \min\limits_{\mathbf{k}} & e_{diff}(\mathbf{y}, t, \mathbf{A}, \mathbf{k}) \\ s.t. & \mathbf{k} \in \mathfrak{K} \end{array} \right. \tag{1.8}$$

$$\left\{ \begin{array}{ll} \min\limits_{\mathbf{k}} & e_{int}(\mathbf{y}, t, \mathbf{A}, \mathbf{k}) \\ s.t. & \mathbf{k} \in \mathfrak{K} \end{array} \right. \tag{1.9}$$

Due to the difficulty of the MINLP, we propose to address first its continuous relaxation, and thus non integer values are obtained. Anyway, the continuous relaxation is multimodal, and the local search used, solved with the nonlinear package `Minos`, [16], is embedded in a multistart procedure.

When the continuous relaxation problem is solved, now we search the integer matrix that best approximates to the continuous one, not just rounding the values (infeasible matrices can appear) but solving a new optimization problem. In this case, the solver `CPLEX`, [8], is used. Since it uses a branch and bound strategy, it gives the global optimum of the problem. Chapter 4 gives more information for this issue.

We have said that the optimization in the differential and integral methods is performed for a fixed matrix $\mathbf{A}$. However, the components of this matrix are also variables in our problem. The method used to find them is a combinatorial method based in the Variable Neighbourhood Search, VNS ([15]). More details about this can be seen in Chapter 5

Finally, Chapter 6 shows the results obtained with the methodology proposed for the data that appear in Appendix A.

Chapter 2

Fitting Concentrations

The goal of this chapter is to compare the behaviour of different models to fit empirical concentrations. Therefore, the problem we want to solve is

$$\min_{\alpha \in \mathcal{A}} \sum_{n=1}^{N_0} (\hat{Y}_{in} - g(\alpha, t_n))^2,$$

where $\hat{Y}_{in}$ are the empirical concentrations of the specie E_i in the time step t_n, $g(\alpha, t)$ is the fitting function, $\alpha \in \mathcal{A}$ are the parameters of the function g and $\mathcal{A}$ is the parameter space.

These curves will be used later to obtain the derivative values needed to evaluate the objective function in the differential method. Let us see the model functions.

2.1 Weibull

We propose a model based on the Weibull distribution (see [3]), the scaled density of which is given by

$$g(t, \alpha) = \alpha_3 \alpha_1 \left(\frac{1}{\alpha_2^{\alpha_1}} \right) (t+1)^{\alpha_1 - 1} e^{-\left(\frac{t+1}{\alpha_2}\right)^{\alpha_1}},$$

where $\alpha_1 > 0$ is the shape parameter, $\alpha_2 > 0$ is the scale parameter and $\alpha_3 > 0$ is a new extra parameter that is used for adapting the data to the function. Note that α_3 is not part of the probability density function.

Also, notice that we consider the translated probability function in order to avoid the problems of division by zero that we can find.

The parameters α_1 and α_2 are related to the mean μ and variance σ^2 of the Weibull distribution through the following relations:

$$\mu = \frac{1}{\alpha_2}\Gamma\left(1+\frac{1}{\alpha_1}\right) \tag{2.1}$$

$$\sigma^2 = \frac{1}{\alpha_2^2}\left(\Gamma\left(1+\frac{2}{\alpha_1}\right)-\Gamma^2\left(1+\frac{1}{\alpha_1}\right)\right)^{1/2}, \tag{2.2}$$

where Γ denotes the Gamma function, given by

$$\Gamma(z)=\int_0^\infty e^{-\tau}\tau^{z-1}d\tau$$

This is a convex function on the interval $0 < z < +\infty$.

According to (2.1), for a fixed α_1, both the mean and the variance of the Weibull distribution decrease with increasing α_2, with a quadratic decrease in the latter case. On the other hand, for a fixed α_2, the mean, μ is a convex-concave function of α_1 with a vertical asymptote at $\alpha_1 = 0$ and a horizontal asymptote given by the line $\mu = 1$. When α_2 is fixed, the variance is a convex function of α_1 with both axes serving as asymptotes.

2.2 Logistic

Another model proposed is based on the logistic distribution. Its scaled density function is

$$g(\alpha,t)=\alpha_3\frac{e^{-\frac{(t-\alpha_2)}{\alpha_1}}}{\alpha_1\left(1+e^{-\frac{(t-\alpha_2)}{\alpha_1}}\right)^2},$$

where $\alpha_1 > 0$ is the scale parameter, $\alpha_2 \in \mathbb{R}$ is the localization parameter and $\alpha_3 > 0$ is a new parameter used to adapt the function to the data.

2.3 Mixture of models

In order to make a best fitting of empirical concentrations, a mixture of Weibull and Logistic density functions is proposed. They are a convex combination of density functions of their respective distributions.

The expression of the Weibull mixture ([3]) is:

$$g(\alpha,t)=\alpha_1\left[\alpha_2\left(\alpha_4\frac{1}{\alpha_3^{\alpha_4}}(t+1)^{\alpha_4-1}e^{-\left(\frac{t+1}{\alpha_3}\right)^{\alpha_4}}\right)+(1-\alpha_2)\left(\alpha_6\frac{1}{\alpha_5^{\alpha_6}}(t+1)^{\alpha_6-1}e^{-\left(\frac{t+1}{\alpha_5}\right)^{\alpha_6}}\right)\right],$$

with $\alpha_1>0$ the parameter used to adapt the data to the function, $\alpha_2\in[0,1]$ the parameter of the convex combination, $\alpha_3,\alpha_5>0$ scale parameters of each of the Weibull, and $\alpha_4,\alpha_6>0$ the shape parameters.

On the other hand, the expression of the Logistic mixture is:

$$g(\alpha,t)=\alpha_3\left[\alpha_6\left(\frac{e^{-\frac{(t-\alpha_2)}{\alpha_1}}}{\alpha_1\left(1+e^{-\frac{(t-\alpha_2)}{\alpha_1}}\right)^2}\right)+(1-\alpha_6)\left(\frac{e^{-\frac{(t-\alpha_5)}{\alpha_4}}}{\alpha_4\left(1+e^{-\frac{(t-\alpha_5)}{\alpha_4}}\right)^2}\right)\right],$$

where $\alpha_1,\alpha_4>0$ are scale parameters, $\alpha_2,\alpha_5\in\mathbb{R}$ are localization parameters, $\alpha_3>0$ is the parameter used to adapt the data and $\alpha_6\in[0,1]$ is the parameter of the convex combination.

2.4 Splines

Another fitting model that has been used is the cubic splines ([11], [20], [9]). They are function of an independent variable, t in our case such that at any point its value is given by a third-degree polynomial in t.

Observe that the polynomial at one point, t_n is not necessarily the same as the polynomial at another point, t_m. The places where the polynomial

changes from one form to another are termed knots, nodes or breakpoints, s_q, $q = 1, \ldots, Q$, where Q is the number of knots. Hence, in each interval, $[s_q, s_{q+1}], q = 1, \ldots, Q-1$ a cubic polynomial with the expression in (2.3) is used:

$$g^q(\alpha^{\mathbf{q}}, t) = \alpha_0^q + \alpha_1^q t + \alpha_2^q t^2 + \alpha_3^q t^3, \tag{2.3}$$

where α^q is the vector of the polynomial coefficients.

The most important property of the spline function is that it is a continuous function, and one or more of its derivatives may also be continuous.

In our examples, the number of the knots are adjusted adaptively depending on the number of time steps we have.

2.5 RBF Gaussians

The idea behind RBF interpolation ([17]) is very simple: imagine that every known point t_n influences its surroundings the same way, according to some assumed functional form $\phi(r)$, the radial basis function, that is a function only for a radial distance $r = |t - t_n|$ from the point. Let us try to approximate the interpolating function everywhere by a linear combination of the ϕ's, centered at all the known points,

$$g(t) = \sum_{n=1}^{N_0} \rho_n \phi(|t - t_n|),$$

where the ρ_n's are some unknown set of weights. The weights are determined by requiring that the interpolation be exact at all the known data points. That is equivalent to solving a set of N_0 linear equations in N_0 unknowns for the ρ_n's:

$$g_j = \sum_{n=1}^{N_0} \rho_n \phi(|t_j - t_n|),$$

For our experiments, the last model proposed for fitting the empirical concentrations is the Radial Basis Function (RBF) with a gaussian base, i.e, taking $\phi(|t - t_n|) = e^{-\alpha(t-t_n)^2}$. Its expression is

$$g(t, \alpha) = \sum_{n=1}^{N_0} \rho_n e^{-\alpha|t-t_n|^2},$$

where ρ_n are the weights considered in the linear combination, and α is a parameter tuned by using a variant of cross-validation known as leave-one-out cross validation ([14], [2]). In this algorithm an optimal value of α is selected by minimizing the least square error for a fit to the data based on an interpolant for which one of the centers was left out.

It is important to notice that there is not a model that is always the best, that is to say, the data approximation will depend on these data. Some examples of this remark can be seen in Figures 2.1 - 2.8.

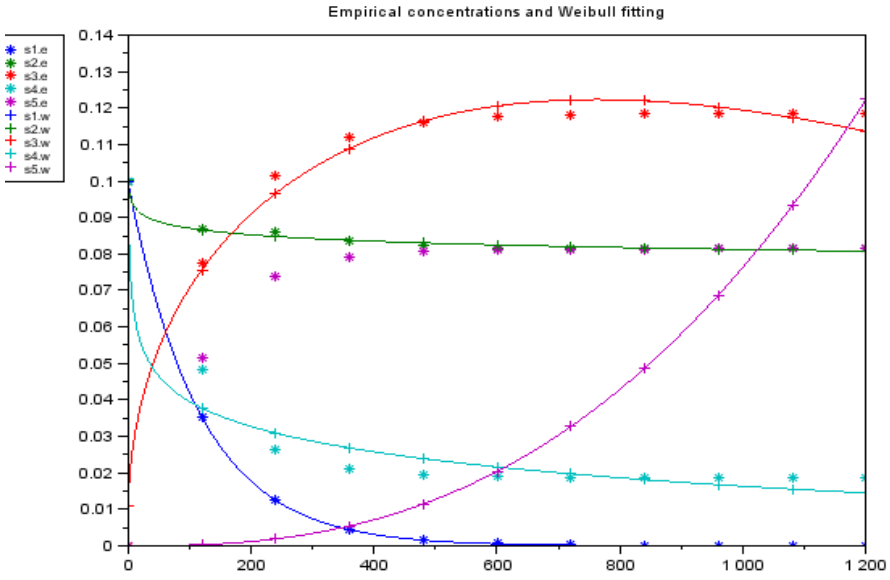

Figure 2.1: Empirical Concentrations and Weibull fitting of the data in [6].

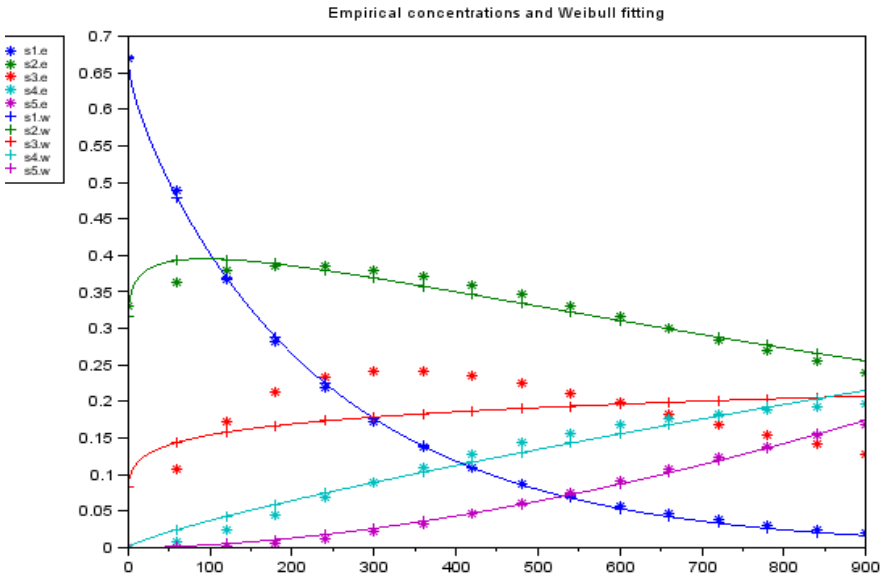

Figure 2.2: Empirical Concentrations and Weibull fitting of the data in [5].

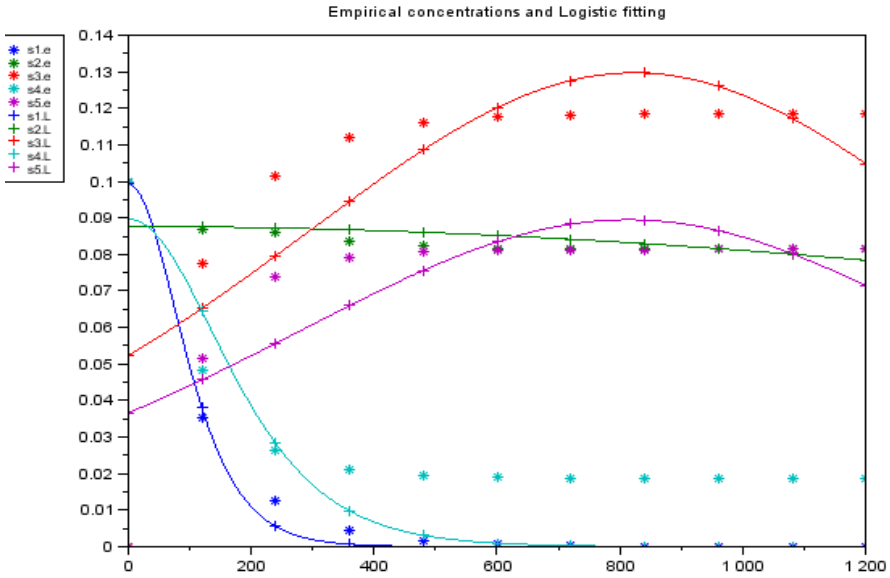

Figure 2.3: Empirical Concentrations and Logistic fitting of the data in [6].

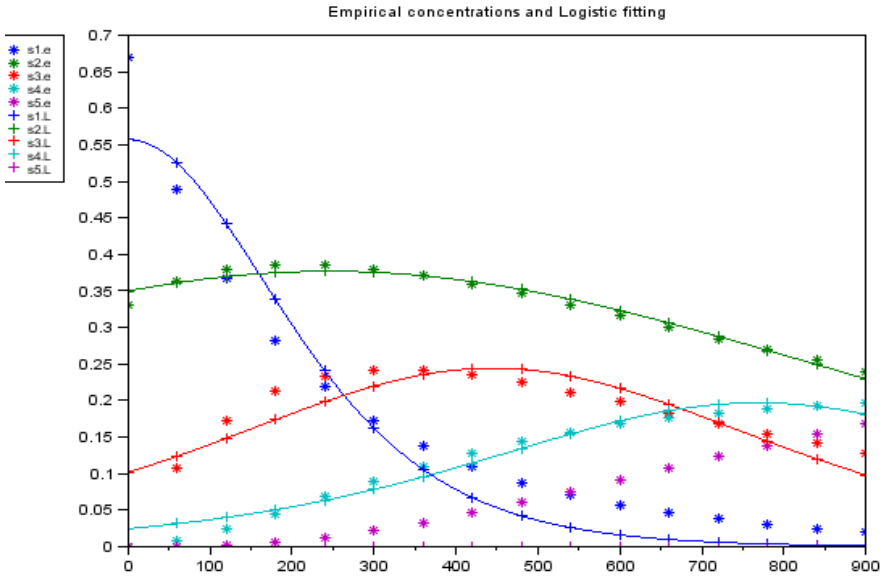

Figure 2.4: Empirical Concentrations and Logistic fitting of the data in [5].

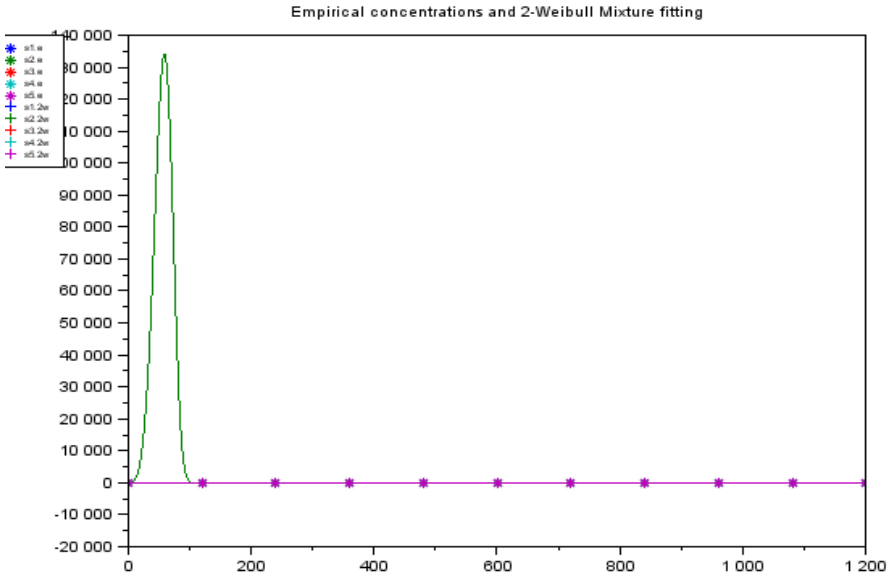

Figure 2.5: Empirical Concentrations and Mixture of 2 Weibull fitting of the data in [6].

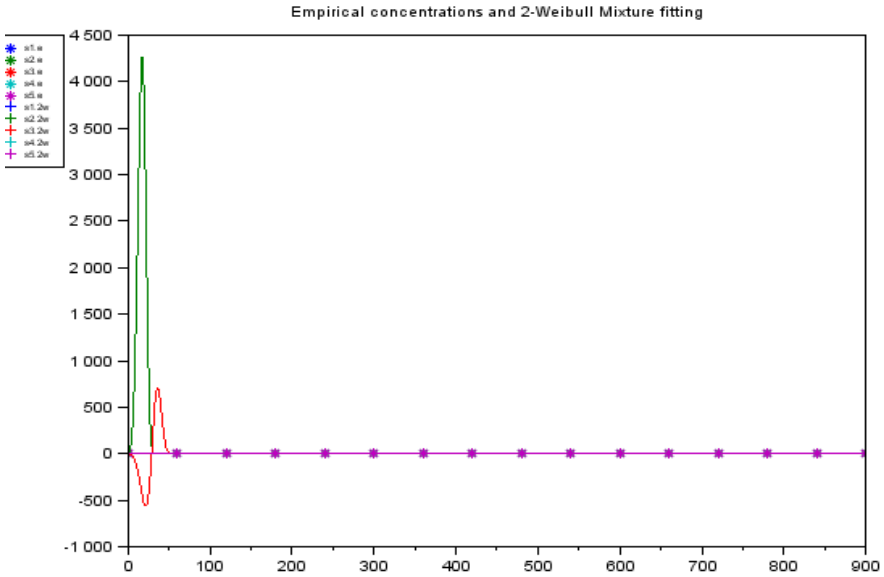

Figure 2.6: Empirical Concentrations and Mixture of 2 Weibull fitting of the data in [5].

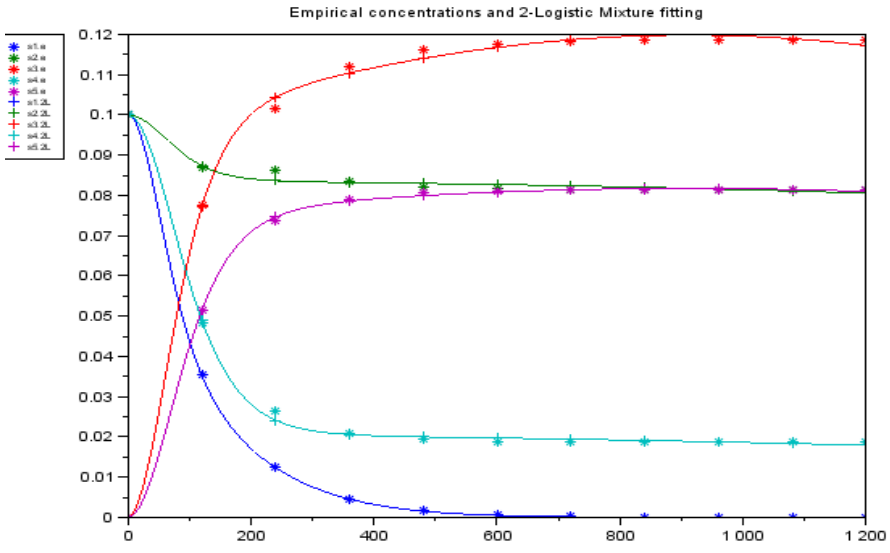

Figure 2.7: Empirical Concentrations and Mixture of 2 Logistic fitting of the data in [6].

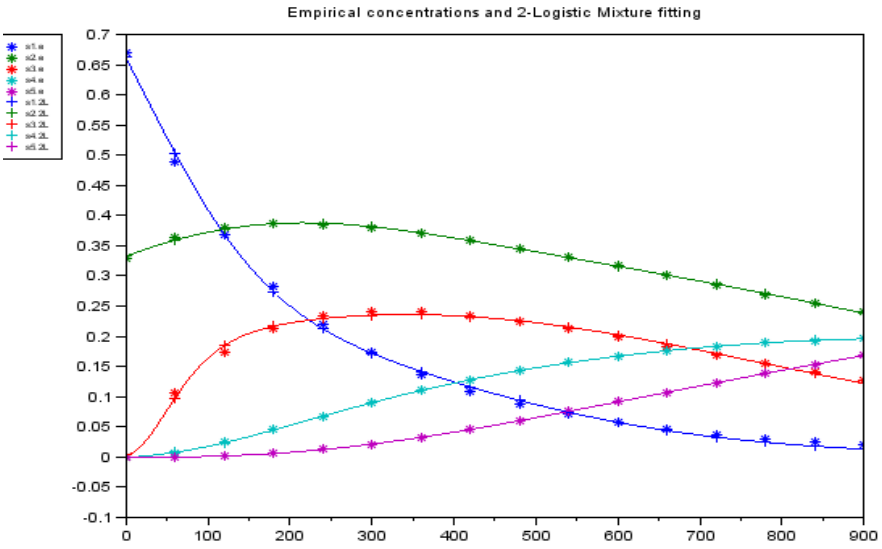

Figure 2.8: Empirical Concentrations and Mixture of 2 Logistic fitting of the data in [5].

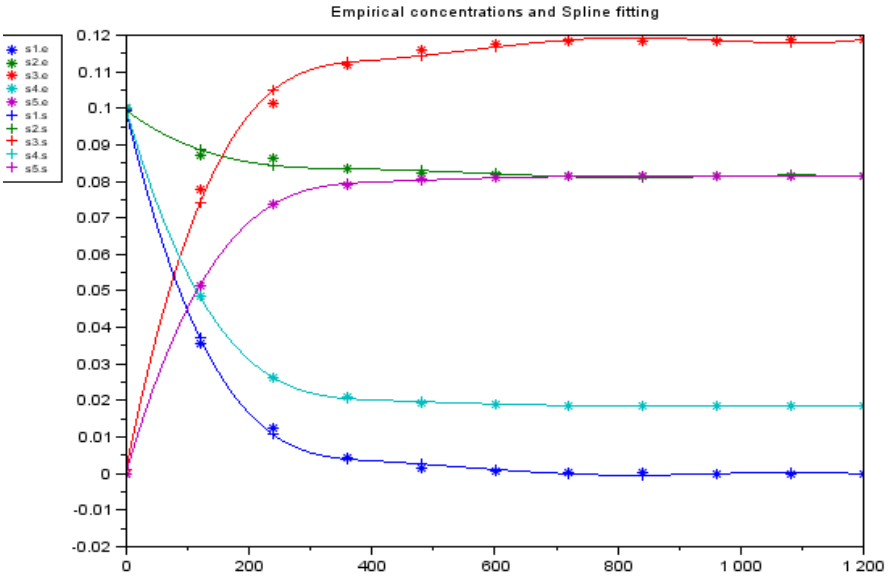

Figure 2.9: Empirical Concentrations and Spline fitting of the data in [6].

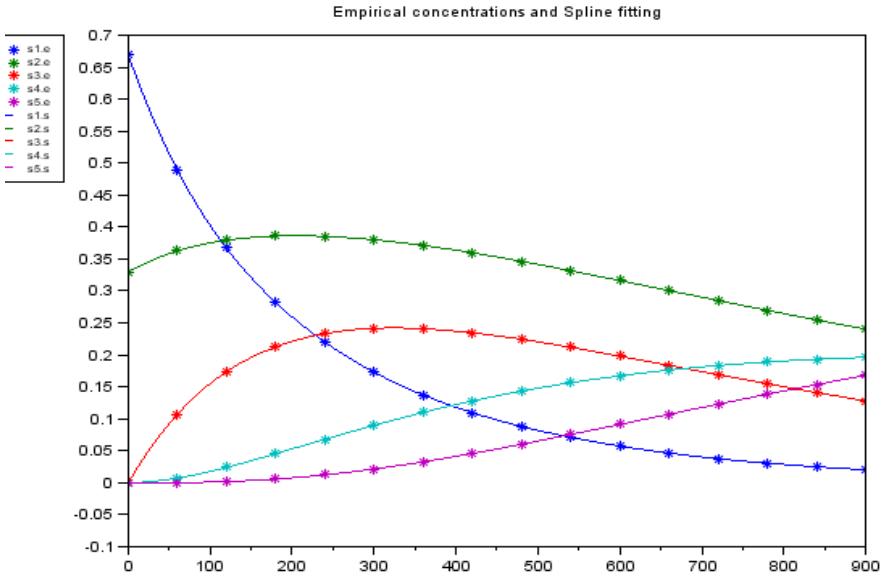

Figure 2.10: Empirical Concentrations and Spline fitting of the data in [5].

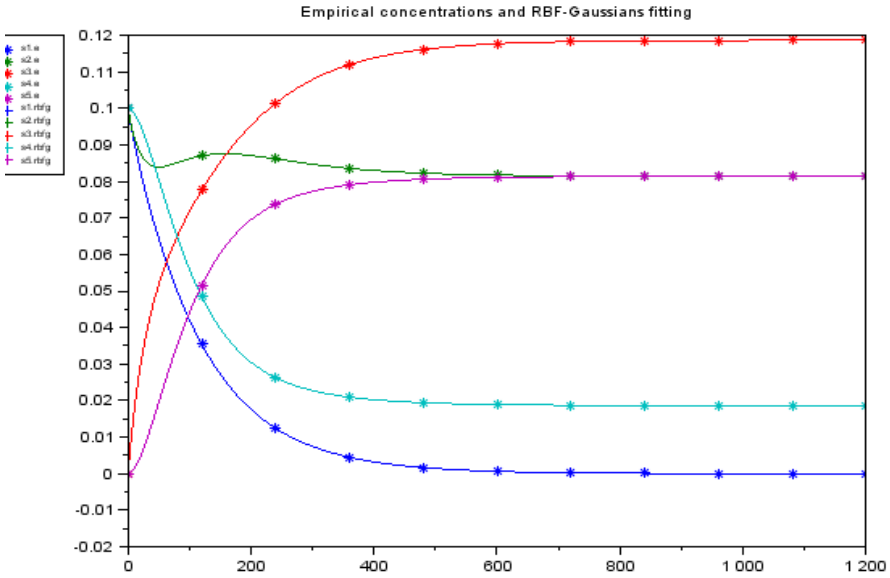

Figure 2.11: Empirical Concentrations and RBF-Gaussians fitting of the data in [6].

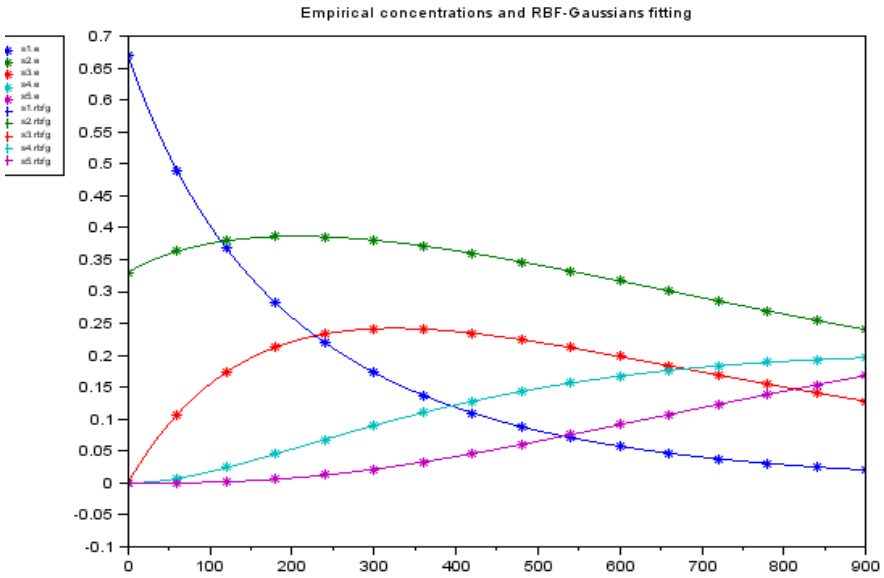

Figure 2.12: Empirical Concentrations and RBF-Gaussians fitting of the data in [5].

Chapter 3

MINLP Formulation

In this chapter we make a detailed description of the formulation of the problem we want to solve. The main objective of this problem is to obtain the chemical reaction model for the concentration dynamics of a set of species that best explains the observed concentrations at the observation instants.

Then, taking as input the species concentrations at some observation instants, we try to identify the integer entries of the stoichiometric matrix, $\mathbf{A}$, (1.1.1) and the rate constants, $\mathbf{k}$.

It is unlikely that a perfect match can be obtained, and therefore one needs to gauge the goodness of fit of the solution found. The empirical concentration, $\hat{y}_{in}$, of specie E_i is known in a set of N_0 time instants. Such $\hat{y}_{in}$ should be close (ideally identical) to the theoretical concentration $y_{in} = y_i(t_n)$, obtained by solving the Cauchy problem expressed by the ODE (1.3) with the initial condition $y(0) = y_0$.

As goodness of fit, the following measure is proposed

$$\sum_{i=1}^{N} \sum_{n=1}^{N_0} w_{in} \left(\hat{y}_{in} - y_{in}\right)^2, \tag{3.1}$$

where coefficients w_{in} are weights reflecting the importance given by the user to errors at each species and time steps.

It is important to note that finding parameters $\mathbf{A}$ and $\mathbf{k}$ by optimizing a goodness of fit measure like (3.1) is a hard challenge, since we face a mixed integer nonlinear programm (MINLP) in which some decision variables take values in integer numbers (the coefficients of $\mathbf{A}$) and some take real values (the rates $\mathbf{k}$).

3.1 Objective Function

Let us see how to formulate the objective function of the problem depending on whether the differential or integral method is used.

3.1.1 Differential Method

The idea of this method is to avoid solving numerically the equation (1.3) as part of the optimization of the goodness of fit function (3.1).

First, a fitting method is used to obtain, from the empirical values $\hat{y}_{in}$ to the full range of positive numbers, $\hat{y}_i(t), t \geq 0$. The procedures tested are described in Chapter 2.

In the differential method, the concentrations are assumed to be given, that is to say, the function $\hat{y}_i(t)$ obtained by the fitting methods are assumed to give an accurate fit to the concentrations functions. In other words,

$$\hat{y}_i(t) \approx y_i(t), \quad \forall t, \forall i = 1, \ldots, N \tag{3.2}$$

As a rough approximation, we assume from (3.2) that the derivatives of the concentrations y are also accurately approximated by the derivatives of the approximations, $\hat{y}$, i.e.,

$$\frac{d}{dt}\hat{y}_i(t_n) \approx \frac{d}{dt}y_i(t_n), \quad \forall i = 1, \ldots, N, \forall n = 1, \ldots, N_0. \tag{3.3}$$

Then, the objective function, $e_{diff}(\mathbf{y}, t, \mathbf{A}, \mathbf{k})$ in (1.8), is used as a surrogate of the goodness of fit criterion (3.1), in which the law of mass action (1.3) is assumed:

$$e_{diff}(\mathbf{y}, t, \mathbf{A}, \mathbf{k}) = \sum_{i=1}^{N}\sum_{n=1}^{N_0} w_{in}\left(\frac{d}{dt}\hat{y}_{in} - \sum_{l=1}^{L}\left(\left(\lambda_i^l - \nu_i^l\right)k_l \prod_{j=1}^{N} y_{jn}^{\nu_j^l}\right)\right)^2 \tag{3.4}$$

Hence, for a fix stoichiometric matrix, $\mathbf{A}$, we try to find the rate reactions values, $\mathbf{k}$, that minimize (3.4).

This method is very fast in terms of computational cost but the results are not very sharp. This is due to the fact that we are assuming that the fitting of the derivative empirical concentration are the same as the derivative of the fitting of the data.

3.1.2 Integral Method

The success of the differential method described in Section 3.1.1 strongly depends on the accuracy of the approximation of the derivative concentrations y by the derivative of the approximations, $\hat{y}$, (3.3).

When this assumption is not true, then the differential method may not yield an accurate solution, but it can be used as an initial solution in a further (more expensive in terms of computational cost) phase, taking as objective function, $e_{int}(\mathbf{y}, t, \mathbf{A}, \mathbf{k})$ in (1.9), the natural goodness of fit (3.1).

The only difference with the differential method is the following: each time the objective function is evaluated, the ODE (1.3), with the initial solution $y(0) = y_0$ must be numerically solved.

3.2 Constraints

Let us explain the different constraints that appear in the optimization problem either with the objective function in (3.2), in the differential method, or (3.1), in the integral method.

We can assume that there are K different chemical elements involved in species, named $H_k, k = 1, \ldots, K$. Then, the formula of specie E_i is

$$E_i = (H_1)_{h_{1i}} \ldots (H_K)_{h_{Ki}}, \quad i = 1, \ldots, N$$

Since atoms are conserved in chemical reactions, we have

$$\sum_{i=1}^{N} h_{ki}\nu_i^l = \sum_{i=1}^{N} h_{ki}\lambda_i^l, \quad k = 1, \ldots, K,\ l = 1, \ldots, L$$

The matrix $(\mathbf{H})_{ki} = h_{ki}$ is called the element matrix. The above equations can be written in a shorter way,

$$\mathbf{HA} = 0 \tag{3.5}$$

From the mass conservation in the l-th chemical reaction, we can easily deduce the second constraint

$$\sum_{i=1}^{N} \mathcal{M}_i \lambda_i^l = \sum_{i=1}^{N} \mathcal{M}_i \nu_i^l, \quad l = 1, \ldots, L$$

It can also be expressed in a matrix form as

$$\mathbf{A}^T \mathcal{M} = 0, \tag{3.6}$$

where $\mathcal{M}$ is the column vector of the molecular masses of species.

The third and fourth constraints below force that, for a given reaction $l = 1, \ldots, L$, the number of molecules of products and reactives, respectively, is bounded:

$$1 \leq \sum_{i=1}^{N} \lambda_i^l \leq C_{max}^{\lambda}, \quad l = 1, \ldots, L \tag{3.7}$$

$$1 \leq \sum_{i=1}^{N} \nu_i^l \leq C_{max}^{\nu}, \quad l = 1, \ldots, L \tag{3.8}$$

The following constraints imply that the stoichiometric coefficients of the products and reactives, respectively must be bounded:

$$0 \leq \lambda_i^l \leq \lambda_{max}, \quad i = 1, \ldots, N, l = 1, \ldots, L \tag{3.9}$$

$$0 \leq \nu_i^l \leq \nu_{max}, \quad i = 1, \ldots, N, l = 1, \ldots, L \tag{3.10}$$

Finally, the constraint

$$k_l \geq 0, \quad l = 1, \ldots, L, \tag{3.11}$$

forces the rate k to be non-negative, and

$$\lambda_i^l, \nu_i^l \in \mathbb{Z}, i = 1, \ldots, N, l = 1, \ldots, L \tag{3.12}$$

forces the stoichiometric matrix to take integer values.

Constraints (3.5)-(3.12) are nice, since they are linear, but they ignore relevant information, namely,

- The minimum number of species, including reactives and products must be not smaller that a given value N_{min}.
- If a species E_i appears in a reaction as a reactive, it cannot appear in the same reaction as a product.
- All reactions are simplified, in the sense that for a fixed reaction l, there exists a species i such that

$$\frac{\lambda_i^l}{\gcd(\lambda_i^l, \nu_i^l, i = 1, \ldots N)} \notin \mathbb{N},$$

  or,

$$\frac{\nu_i^l}{\gcd(\lambda_i^l, \nu_i^l, i = 1, \ldots N)} \notin \mathbb{N},$$

  where *gcd* denotes the great common divisor.

- All species must appear in at least one reaction.

In order to accommodate such constraint, the problem formulation must be enriched by adding new variables and constraints. Indeed, new binary variables are define:

$$z_{ilm}^{\nu} = \begin{cases} 1, & \text{if } \nu_i^l = m \\ 0, & \text{otherwise} \end{cases}, \quad m = 1, \ldots, \nu_{max}$$

and,

$$z_{ilm}^{\lambda} = \begin{cases} 1, & \text{if } \lambda_i^l = m \\ 0, & \text{otherwise} \end{cases}, \quad m = 1, \ldots, \lambda_{max} \tag{3.13}$$

With these new variables we can rewrite products:

$$\lambda_i^l = \sum_{m=1}^{\lambda_{max}} m z_{ilm}^{\lambda}, \quad l = 1, \ldots, L,\ i = 1, \ldots, N \tag{3.14}$$

and, analogously, the reactives

$$\nu_i^l = \sum_{m=1}^{\nu_{max}} m z_{ilm}^{\nu}, \quad l = 1, \ldots, L,\ i = 1, \ldots, N \tag{3.15}$$

With these variables, the condition expressing that the minimum number of species must be not smaller than a given parameter $N_{\min}$ is expressed as

$$\sum_{i=1}^{N} \left(\sum_{m=1}^{\nu_{max}} z_{ilm}^{\nu} + \sum_{m=1}^{\lambda_{max}} z_{ilm}^{\lambda} \right) \geq N_{\min}, \quad l = 1, \ldots, L \tag{3.16}$$

Furthermore, expressing that the same species cannot appear as a reactive and as a product in the same reaction is expressed as

$$\sum_{m=1}^{\nu_{max}} z_{ilm}^{\nu} + \sum_{m=1}^{\lambda_{max}} z_{ilm}^{\lambda} \leq 1, \quad i = 1, \ldots, N,\ l = 1, \ldots, L \tag{3.17}$$

Expressing that reactions are already simplified is written as

$$\sum_{i=1}^{N} \left(z_{il1}^{\lambda} + z_{il1}^{\nu} \right) \geq 1, \quad l = 1, \ldots, L \tag{3.18}$$

Finally, expressing that each species appears in at least one reaction is easily shown to be written as

$$\sum_{l=1}^{L} \left(\sum_{m=1}^{\nu_{max}} z_{ilm}^{\nu} + \sum_{m=1}^{\lambda_{max}} z_{ilm}^{\lambda} \right) \geq 1, \quad i = 1, \ldots, N \tag{3.19}$$

Summarizing, the optimization problems needed to be solved are (3.20) in the differential method and (3.21) in the integral method.

(3.20)

$$
\begin{cases}
\min\limits_{k_l,\lambda_i^l,\nu_i^l,z_{ilm}^{\lambda},z_{ilm}^{\nu}} & \sum\limits_{i=1}^{N}\sum\limits_{n=1}^{N_0} w_{in}\left(\dfrac{d}{dt}\hat{y}_{in}-\sum\limits_{l=1}^{L}\left(\left(\lambda_i^l-\nu_i^l\right)k_l\prod\limits_{j=1}^{N}y_{jn}^{\nu_j^l}\right)\right)^2 \\
\text{s.t.} & \mathbf{H}(\mathbf{P}-\mathbf{R})=0 \\
& (\mathbf{P}-\mathbf{R})^T\mathcal{M}=0 \\
& 1\le\sum\limits_{i=1}^{N}\lambda_i^l\le C_{max}^{\lambda},\quad l=1,\ldots,L \\
& 1\le\sum\limits_{i=1}^{N}\nu_i^l\le C_{max}^{\nu},\quad l=1,\ldots,L \\
& \lambda_i^l=\sum\limits_{m=1}^{\lambda_{max}}mz_{ilm}^{\lambda},\quad l=1,\ldots,L,\ i=1,\ldots,N \\
& \nu_i^l=\sum\limits_{m=1}^{\nu_{max}}mz_{ilm}^{\nu},\quad l=1,\ldots,L,\ i=1,\ldots,N \\
& \sum\limits_{i=1}^{N}\left(\sum\limits_{m=1}^{\nu_{max}}z_{ilm}^{\nu}+\sum\limits_{m=1}^{\lambda_{max}}z_{ilm}^{\lambda}\right)\ge N_{\min},\quad l=1,\ldots,L \\
& \sum\limits_{m=1}^{\nu_{max}}z_{ilm}^{\nu}+\sum\limits_{m=1}^{\lambda_{max}}z_{ilm}^{\lambda}\le 1,\quad i=1,\ldots,N,\ l=1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(z_{il1}^{\lambda}+z_{il1}^{\nu}\right)\ge 1,\quad l=1,\ldots,L \\
& \sum\limits_{l=1}^{L}\left(\sum\limits_{m=1}^{\nu_{max}}z_{ilm}^{\nu}+\sum\limits_{m=1}^{\lambda_{max}}z_{ilm}^{\lambda}\right)\ge 1,\quad i=1,\ldots,N \\
& 0\le\lambda_i^l\le\lambda_{max},\quad i=1,\ldots,N,l=1,\ldots,L \\
& 0\le\nu_i^l\le\nu_{max},\quad i=1,\ldots,N,l=1,\ldots,L \\
& k_l\ge 0,\quad l=1,\ldots,L \\
& \lambda_i^l,\nu_i^l\in\mathbb{Z},i=1,\ldots,N,l=1,\ldots,L \\
& z_{ilm}^{\nu}\in\{0,1\},\quad i=1,\ldots,N,\ m=1,\ldots,\nu_{max},\ l=1,\ldots,L \\
& z_{ilm}^{\lambda}\in\{0,1\},\quad i=1,\ldots,N,\ m=1,\ldots,\lambda_{max},\ l=1,\ldots,L
\end{cases}
$$

(3.21)

$$
\begin{cases}
\min\limits_{k_l,\lambda_i^l,\nu_i^l,z_{ilm}^{\lambda},z_{ilm}^{\nu}} & \sum\limits_{i=1}^{N}\sum\limits_{n=1}^{N_0} w_{in}\left(\hat{y}_{in}-y_{in}\right)^2 \\
\text{s.t.} & \mathbf{H}(\mathbf{P}-\mathbf{R})=0 \\
& (\mathbf{P}-\mathbf{R})^T\mathcal{M}=0 \\
& 1\le\sum\limits_{i=1}^{N}\lambda_i^l\le C_{max}^{\lambda},\quad l=1,\ldots,L \\
& 1\le\sum\limits_{i=1}^{N}\nu_i^l\le C_{max}^{\nu},\quad l=1,\ldots,L \\
& \lambda_i^l=\sum\limits_{m=1}^{\lambda_{max}} m z_{ilm}^{\lambda},\quad l=1,\ldots,L,\ i=1,\ldots,N \\
& \nu_i^l=\sum\limits_{m=1}^{\nu_{max}} m z_{ilm}^{\nu},\quad l=1,\ldots,L,\ i=1,\ldots,N \\
& \sum\limits_{i=1}^{N}\left(\sum\limits_{m=1}^{\nu_{max}} z_{ilm}^{\nu}+\sum\limits_{m=1}^{\lambda_{max}} z_{ilm}^{\lambda}\right)\ge N_{\min},\quad l=1,\ldots,L \\
& \sum\limits_{m=1}^{\nu_{max}} z_{ilm}^{\nu}+\sum\limits_{m=1}^{\lambda_{max}} z_{ilm}^{\lambda}\le 1,\quad i=1,\ldots,N,\ l=1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(z_{il1}^{\lambda}+z_{il1}^{\nu}\right)\ge 1,\quad l=1,\ldots,L \\
& \sum\limits_{l=1}^{L}\left(\sum\limits_{m=1}^{\nu_{max}} z_{ilm}^{\nu}+\sum\limits_{m=1}^{\lambda_{max}} z_{ilm}^{\lambda}\right)\ge 1,\quad i=1,\ldots,N \\
& 0\le\lambda_i^l\le\lambda_{max},\quad i=1,\ldots,N,l=1,\ldots,L \\
& 0\le\nu_i^l\le\nu_{max},\quad i=1,\ldots,N,l=1,\ldots,L \\
& k_l\ge 0,\quad l=1,\ldots,L \\
& \lambda_i^l,\nu_i^l\in\mathbb{Z},i=1,\ldots,N,l=1,\ldots,L \\
& z_{ilm}^{\nu}\in\{0,1\},\quad i=1,\ldots,N,\ m=1,\ldots,\nu_{max},\ l=1,\ldots,L \\
& z_{ilm}^{\lambda}\in\{0,1\},\quad i=1,\ldots,N,\ m=1,\ldots,\lambda_{max},\ l=1,\ldots,L
\end{cases}
$$

Observe that equations (3.14) and (3.15) allow us to express (3.20) and (3.21) in a simplest form, where the variables are $z^{\lambda}_{ilm}, z^{\nu}_{ilm}, k_l$. It can be seen in (3.22) and (3.23), respectively.

(3.22)

$$
\begin{cases}
\min\limits_{k_l, z^{\lambda}_{ilm}, z^{\nu}_{ilm}} & \sum\limits_{i=1}^{N}\sum\limits_{n=1}^{N_0} w_{in}\left(\dfrac{d}{dt}\hat{y}_{in} - \sum\limits_{l=1}^{L}\left(\left(\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm} - \sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm}\right) k_l \prod\limits_{j=1}^{N} y_{jn}^{\left(\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{jlm}\right)}\right)\right)^2 \\
\text{s.t.} & \sum\limits_{i=1}^{N}\left(h_{ki}\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm}\right) = \sum\limits_{i=1}^{N}\left(h_{ki}\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm}\right), \quad k = 1,\ldots,K,\ l = 1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(\mathcal{M}_i\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm}\right) = \sum\limits_{i=1}^{N}\left(\mathcal{M}_i\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm}\right), \quad l = 1,\ldots,L \\
& 1 \leq \sum\limits_{i=1}^{N}\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm} \leq C^{\lambda}_{max}, \quad l = 1,\ldots,L \\
& 1 \leq \sum\limits_{i=1}^{N}\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm} \leq C^{\nu}_{max}, \quad l = 1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(\sum\limits_{m=1}^{\nu_{max}} z^{\nu}_{ilm} + \sum\limits_{m=1}^{\lambda_{max}} z^{\lambda}_{ilm}\right) \geq N_{\min}, \quad l = 1,\ldots,L \\
& \sum\limits_{m=1}^{\nu_{max}} z^{\nu}_{ilm} + \sum\limits_{m=1}^{\lambda_{max}} z^{\lambda}_{ilm} \leq 1, \quad i = 1,\ldots,N,\ l = 1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(z^{\lambda}_{il1} + z^{\nu}_{il1}\right) \geq 1, \quad l = 1,\ldots,L \\
& \sum\limits_{l=1}^{L}\left(\sum\limits_{m=1}^{\nu_{max}} z^{\nu}_{ilm} + \sum\limits_{m=1}^{\lambda_{max}} z^{\lambda}_{ilm}\right) \geq 1, \quad i = 1,\ldots,N \\
& k_l \geq 0, \quad l = 1,\ldots,L \\
& z^{\nu}_{ilm} \in \{0,1\}, \quad i = 1,\ldots,N,\ m = 1,\ldots,\nu_{max},\ l = 1,\ldots,L \\
& z^{\lambda}_{ilm} \in \{0,1\}, \quad i = 1,\ldots,N,\ m = 1,\ldots,\lambda_{max},\ l = 1,\ldots,L
\end{cases}
$$

(3.23)

$$
\begin{cases}
\min\limits_{k_l, z^{\lambda}_{ilm}, z^{\nu}_{ilm}} & \sum\limits_{i=1}^{N}\sum\limits_{n=1}^{N_0} w_{in}\left(\hat{y}_{in}-y_{in}\right)^2 \\
\text{s.t.} & \sum\limits_{i=1}^{N}\left(h_{ki}\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm}\right)=\sum\limits_{i=1}^{N}\left(h_{ki}\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm}\right), \quad k=1,\ldots,K,\, l=1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(\mathcal{M}_i\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm}\right)=\sum\limits_{i=1}^{N}\left(\mathcal{M}_i\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm}\right), \quad l=1,\ldots,L \\
& 1\le\sum\limits_{i=1}^{N}\sum\limits_{m=1}^{\lambda_{max}} m z^{\lambda}_{ilm}\le C^{\lambda}_{max}, \quad l=1,\ldots,L \\
& 1\le\sum\limits_{i=1}^{N}\sum\limits_{m=1}^{\nu_{max}} m z^{\nu}_{ilm}\le C^{\nu}_{max}, \quad l=1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(\sum\limits_{m=1}^{\nu_{max}} z^{\nu}_{ilm}+\sum\limits_{m=1}^{\lambda_{max}} z^{\lambda}_{ilm}\right)\ge N_{\min}, \quad l=1,\ldots,L \\
& \sum\limits_{m=1}^{\nu_{max}} z^{\nu}_{ilm}+\sum\limits_{m=1}^{\lambda_{max}} z^{\lambda}_{ilm}\le 1, \quad i=1,\ldots,N,\, l=1,\ldots,L \\
& \sum\limits_{i=1}^{N}\left(z^{\lambda}_{il1}+z^{\nu}_{il1}\right)\ge 1, \quad l=1,\ldots,L \\
& \sum\limits_{l=1}^{L}\left(\sum\limits_{m=1}^{\nu_{max}} z^{\nu}_{ilm}+\sum\limits_{m=1}^{\lambda_{max}} z^{\lambda}_{ilm}\right)\ge 1, \quad i=1,\ldots,N \\
& k_l\ge 0, \quad l=1,\ldots,L \\
& z^{\nu}_{ilm}\in\{0,1\}, \quad i=1,\ldots,N,\, m=1,\ldots,\nu_{max},\, l=1,\ldots,L \\
& z^{\lambda}_{ilm}\in\{0,1\}, \quad i=1,\ldots,N,\, m=1,\ldots,\lambda_{max},\, l=1,\ldots,L
\end{cases}
$$

Chapter 4

Continuous Relaxation and Projection over the Integers

The first step in the search of the optimal stoichiometric matrix and reaction rates of the problem (3.20) is to obtain reasonably good starting values of matrix $\mathbf{A}$. This process has different phases to properly handle the fact that some of the variables take integer values.

The first phase is to find an approximation to $\mathbf{A}$ in real numbers. That is simply done by omitting the integrity constraints; in other words, handling a purely continuous nonlinear optimization problem in which the λ variables are in the interval $[0, \lambda_{max}]$, ν variables in $[0, \nu_{max}]$ and z in $[0, 1]$. This problem will be called the continuous relaxation of (3.20).

It has been observed that the problem is highly multimodal, and thus standard local search solvers, as the one used in our numerical experiments, `MINOS` [16], may get trapped at local optima. Therefore a multistart technique is performed: several runs of the same problem are made but changing the starting point randomly selected.

For each run r, a pair $(\mathbf{A}^r, \mathbf{k}^r)$ is obtained. Since $\mathbf{A}^r$ is a continuous matrix (it is unlikely its components were integer numbers) a second phase, projection over the integers, is made. This projection cannot be made by simply rounding the values, because infeasible matrices will appear due to the many constraints in the problem. The method used to find the feasible matrix that best fits the continuous one is by solving a new quadratic convex optimization problem with linear constraints.

$$\begin{cases} \min\limits_{\lambda_i^l,\nu_i^l,z_{ilm}^{\lambda},z_{ilm}^{\nu}} & \|(\mathbf{P}-\mathbf{R})-\mathbf{A}^r\|^2 \\ s.t. & \text{constraints in (3.20)} \\ & \mathbf{k}=\mathbf{k}^r \end{cases}$$

In this problem, the variables are the integer stoichiometric coefficients of the matrices $\mathbf{P}$ and $\mathbf{R}$. For the resolution of this problem the global optimizer `CPLEX`, [8], is used.

This method is going to be applied in some different experiments where for example, the weights w_{in} in the continuous relaxation of (3.20) are changed or the problem is solved with the formulation in z_{ilm}^{ν} and z_{ilm}^{ν}, or with z_{ilm}^{ν},z_{ilm}^{ν}, λ_i^l and ν_i^l.

As we have said before, this method is applied for determined experiments and in each experiment this methodology is applied for a fixed number of runs. Once we have all these integer matrices the optimization problem of differential method is applied, (3.20), keeping the best matrix of each experiment in terms of objective value. With these matrices, now we apply to the integral method optimization problem, (3.21). The one that gives the least value in the objective function will be taken as an initial stoichiometric matrix.

A flow diagram of process explained above can be seen in Figure 4.1.

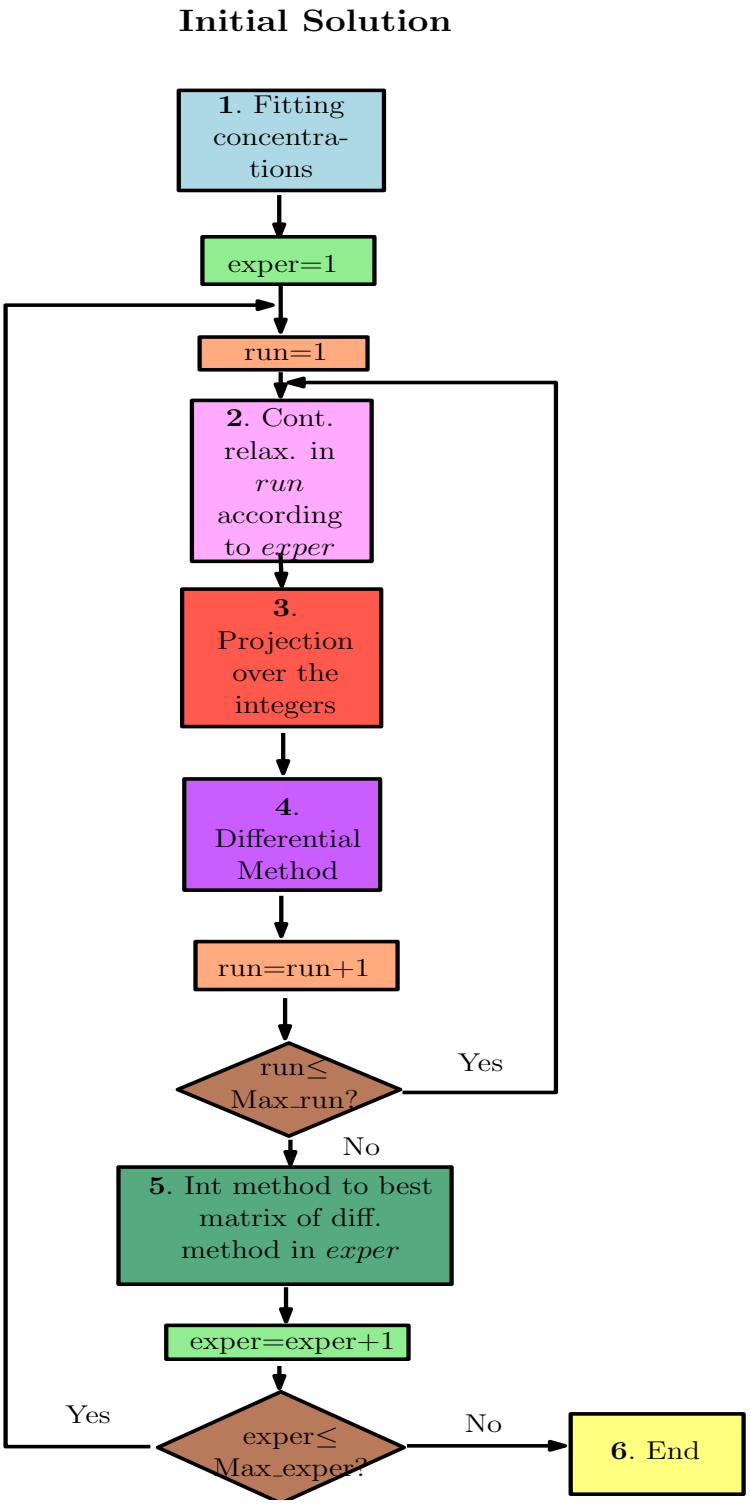

Figure 4.1: Flow diagram of the initial solution procedure.

Chapter 5

Combinatorial Method

In Chapter 3 it is explained that, for a fixed stoichiometric matrix, $\mathbf{A}$, the differential and the integral method can be applied. However, $\mathbf{A}$ is also a decision variable, and thus further efforts are needed to explore the space of stoichiometric matrices.

As initial matrix, the one obtained as in the process scheme explained in Chapter 4 is used. The following methodology is an adaptation to this problem of the Variable Neighbourhood Search algorithm method, VNS, [15].

VNS is a technique used in combinatorial optimization problems, such as the Travel Salesman Problem (TSP), [15]. The idea is starting from an initial solution and improve the objective value changing, what the authors call, the neighborhood in the search. Contrary to most other local search method VNS does not follow a trajectory, but explores increasingly distant neighbourhoods of the current incumbent solution, and jumps from there to a new one if and only if an improvement was made.

A pseudocode of VNS ([13]) is given in Figure 5.1.

In our problem a perturbation $\tilde{\mathbf{A}}$ of the given solution $\mathbf{A}$ is generated, and then the differential method, (3.20) is applied to obtain $\tilde{\mathbf{k}}$. If $(\tilde{\mathbf{A}}, \tilde{\mathbf{k}})$ gives a better objective value, the process is repeated starting from such solution. Otherwise, a new perturbation is generated. This process continues until a stopping rule, related with number of iterations or running time is reached.

After such stopping rule is fullfilled, if we want more accuracy in the results, the process can be repeated optimizing using the integral method, (3.21), instead of the differential one.

Initialization. Select the set of neighborhood structures $\mathcal{N}_k$, $k = 1, \ldots, k_{max}$, that will be used in the search; find an initial solution x; choose a stopping condition;

Repeat the following until the stopping condition is met:
(1) Set $k \leftarrow 1$; (2) Until $k = k_{max}$, repeat the following steps:
(a) *Shaking.* Generate a point x' at random from the k^{th} neighborhood of x ($x' \in \mathcal{N}_k(x)$);
(b) *Local search.* Apply some local search method with x' as initial solution; denote with x'' the so obtained local optimum;
(c) *Move or not.* If this local optimum is better than the incumbent, move there ($x \leftarrow x''$), and continue the search with $\mathcal{N}_1$ ($k \leftarrow 1$); otherwise, set $k \leftarrow k + 1$;

Fig. 1. Steps of the basic VNS.

Figure 5.1: Pseudocode of VNS process.

The key issue, as in all applications of VNS, is how to design the perturbations (neighbourhoods) of a given stoichiometric matrix. We define a perturbation of radius l as follows: l reactions from $\mathbf{A}$ are chosen at random and replaced by l feasible reactions from a list previously calculated.

To build this list where all the feasible reactions are, we first choose the number of nonzero species, $1, \ldots, C^{\nu}_{max}$ that will play as reactives. For each choice, there is a combination of $1, \ldots, C^{\lambda}_{max}$ products. It is important to observe that the species in each side must be different and the stoichiometric components are in the set $\{0, \ldots, \nu_{max}\}$ for the reactives and $\{0, \ldots, \lambda_{max}\}$ for the products.

In the cases, where the molecular weight vector, $\mathcal{M}$ and/or the elemental matrix, $\mathbf{H}$ are known, the feasible reactions must satisfy the constraints (3.6) and (3.5) respectively. Otherwise, a larger list is build.

Let us show some examples of the VNS process in our particular cases. For the data set in A.1, where the matrix $\mathbf{H}$ and the vector $\mathcal{M}$ are known, the list of feasible reactions is not too large, as it can be seen in Table 5.1. In this case the values $\lambda_{max} = \nu_{max} = 2$ and $C^{\lambda}_{max} = C^{\nu}_{max} = 3$.

1	-1	-1	-1	0	1	0	1
1	-1	0	1	-1	-1	1	0
-2	2	1	0	1	0	-1	-1
0	0	-1	-2	1	2	-1	1
0	0	1	2	-1	-2	1	-1

Table 5.1: Feasible reactions of the data set A.1.

$$\mathbf{A}_{ini} = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix} \tag{5.1}$$

With this matrix above, (5.1), obtained as Chapter 4 explains (more details of this matrix can be seen in Chapter 6), we can make a perturbation of radius 1. A column of $\mathbf{A}_{ini}$ is chosen randomly, for instance we can assume, that it is the first one, i.e.:

$$\mathbf{A}_{ini} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

Then we replace this column by one of the list in Table 5.1, for example, the last one:

$$\mathbf{A}_{ini} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

Then, the perturbed stoichiometric matrix is the one that appears in (5.2)

$$\mathbf{A}_{ini} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \tag{5.2}$$

Let us now explain the pseudocodes used to optimize the stoichiometric matrix $\mathbf{A}$ and the rate coefficient vector $\mathbf{k}$ with the differential or the integral method.

ALGORITHM 1

1) Begin with an initial matrix $\mathbf{A}_{ini}$.
2) Optimize the variables in $\mathbf{k}$ with the differential method and obtain an initial objective value, obj_{ini}.
3) p:=1. Set $\mathbf{A}_{opt} := \mathbf{A}_{ini}$ and $obj_{opt} := obj_{ini}$.
4) Until $p = p_{max}$, perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.

  4.1)Perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.
  4.2) Fixed $\mathbf{A}$, optimize $\mathbf{k}$ with the differential method.
      The objective value is obj.
  4.3) If $obj < obj_{opt}$, then set $\mathbf{A}_{opt} := \mathbf{A}, obj_{opt} := obj$.
  4.4) Let $p = p + 1$.

5) Optimize $\mathbf{k}$ with $\mathbf{A}_{opt}$ fixed and the differential method.

Table 5.2: Algorithm 1

ALGORITHM 2

1) Begin with an initial matrix $\mathbf{A}_{ini}$.
2) Optimize the variables in $\mathbf{k}$ with the differential method and obtain an initial objective value, obj_{ini}.
3) p:=1. Set $\mathbf{A}_{opt} := \mathbf{A}_{ini}$ and $obj_{opt} := obj_{ini}$.
4) Until $p = p_{max}$, perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.

4.1)Perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.
4.2) Fixed $\mathbf{A}$, optimize $\mathbf{k}$ with the differential method.
The objective value is obj.
4.3) If $obj < obj_{opt}$, then set $\mathbf{A}_{opt} := \mathbf{A}, obj_{opt} := obj$.
4.4) Let $p = p + 1$.

5) Optimize $\mathbf{k}$ with $\mathbf{A}_{opt}$ fixed and the integral method.

Table 5.3: Algorithm 2

ALGORITHM 3

1) Begin with an initial matrix $\mathbf{A}_{ini}$.
2) Optimize the variables in $\mathbf{k}$ with the integral method and obtain an initial objective value, obj_{ini}.
3) p:=1. Set $\mathbf{A}_{opt} := \mathbf{A}_{ini}$ and $obj_{opt} := obj_{ini}$.
4) Until $p = p_{max}$, perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.

4.1)Perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.
4.2) Fixed $\mathbf{A}$, optimize $\mathbf{k}$ with the integral method.
The objective value is obj.
4.3) If $obj < obj_{opt}$, then set $\mathbf{A}_{opt} := \mathbf{A}, obj_{opt} := obj$.
4.4) Let $p = p + 1$.

5) Optimize $\mathbf{k}$ with $\mathbf{A}_{opt}$ fixed and the differential method.

Table 5.4: Algorithm 3

ALGORITHM 4

1) Begin with an initial matrix $\mathbf{A}_{ini}$.
2) Optimize the variables in $\mathbf{k}$ with the integral method and obtain an initial objective value, obj_{ini}.
3) p:=1. Set $\mathbf{A}_{opt} := \mathbf{A}_{ini}$ and $obj_{opt} := obj_{ini}$.
4) Until $p = p_{max}$, perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.

4.1)Perturb $\mathbf{A}_{opt}$ and obtain $\mathbf{A}$.
4.2) Fixed $\mathbf{A}$, optimize $\mathbf{k}$ with the integral method.
The objective value is obj.
4.3) If $obj < obj_{opt}$, then set $\mathbf{A}_{opt} := \mathbf{A}, obj_{opt} := obj$.
4.4) Let $p = p + 1$.

5) Optimize $\mathbf{k}$ with $\mathbf{A}_{opt}$ fixed and the integral method.

Table 5.5: Algorithm 4

Chapter 6

Examples

In this chapter, results obtained by applying the procedure explained in Chapters 3 and 5 from the data in Appendix A are shown. It is important to note that in both data sets the differential method has been applied using the derivatives of the spline fittings. Observe that the column order in the stoichiometric matrix does not affect the solution, since the columns represent the number of the reaction and they can be sorted in the order we choose, keeping in mind that this order must correspond to the components of $\mathbf{k}$.

The maximum time allowed for the experiments in the combinatory using the differential method, first, and, then, the integral one has been 6 hours in each one, that is to say the programme has been running for 12 hours. For the experiments, the solver `Scilab`, [7], has been used.

Furthermore, the parameters $C^{\lambda}_{max}, C^{\nu}_{max}, \lambda_{max}, \nu_{max}, N_{min}$ that appear in (3.20) and (3.21) are taken as

$$C^{\lambda}_{max} = C^{\nu}_{max} = 3$$
$$\lambda_{max} = \nu_{max} = 2$$

Also,

$$N_{min} = \begin{cases} 3, & \text{in } A.1 \\ 2, & \text{in } A.2 \end{cases}$$

6.1 Data Set 1

From the data of A.1, the initial stoichiometric matrix obtained, as explained in Chapter 4 and the **k** reaction rate vector obtained with this matrix, applying the differential method and integral method with their respective objective values are in Table 6.1.

Initial stoichiometric matrix	$\mathbf{A}_{ini} = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix}$
Int. method rate vector	$\mathbf{k}_{ini-diff} = \begin{pmatrix} 0.1693065088 \\ 0.0000100000 \\ 0.0192887246 \end{pmatrix}$
Diff. obj value	$2.0493407576 \cdot 10^{-12}$
Diff. method rate vector	$\mathbf{k}_{ini-int} = \begin{pmatrix} 0.1068719346 \\ 0.0000100000 \\ 0.0188922986 \end{pmatrix}$
Int. obj value	$5.0654641711 \cdot 10^{-8}$

Table 6.1: Initial stoichiometric matrix and reaction vector for the differential and integral method in data set A.1.

Note, that the initial matrix has captured one of the three correct reactions.

The plots of the concentrations obtained with the matrix in Table 6.1 and the values of the parameter vector **k** with the differential and the integral method (see also Table 6.1) are shown in Figures 6.1 and 6.2 respectively.

The next step of the procedure proposed is the combinatorial method just using the differential method (Chapter 5). The matrix **A** obtained at the end of this step is

$$\mathbf{A}_{comb-diff} = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & -1 \\ 2 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \tag{6.1}$$

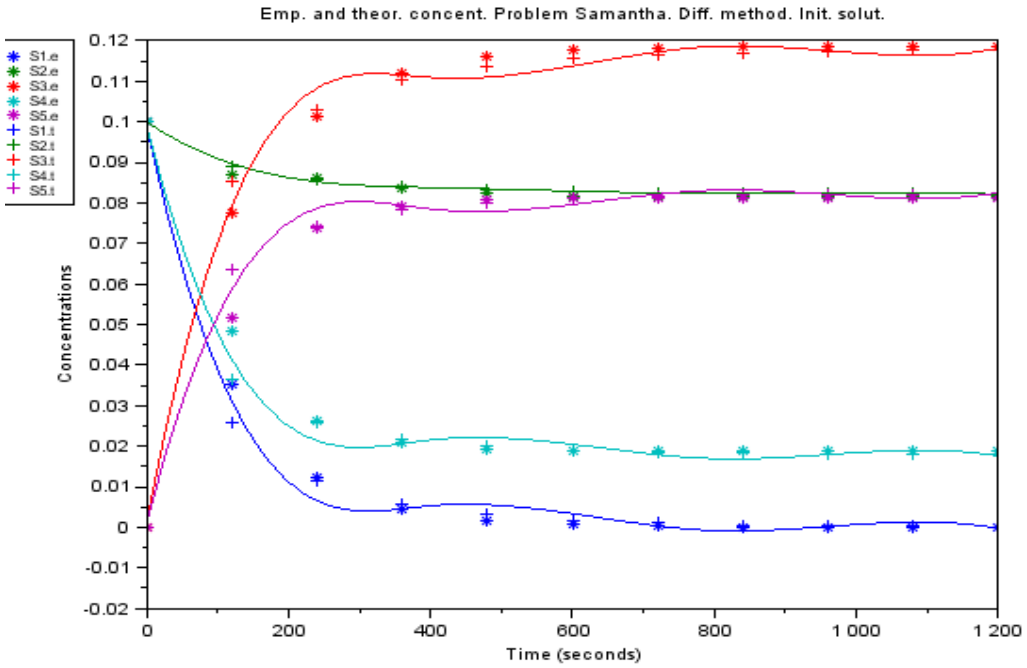

Figure 6.1: Empirical and theoretical concentrations of the data in [6] with the initial parameters and the differential method.

The rate vector, $\mathbf{k}$ calculated using matrix in (6.1) with the Algorithms 1 and 2 (see Tables 5.2 and 5.3 respectively) are written with their objectives values in Table 6.2.

Let us note that in this case, in which the molecular weight vector $\mathcal{M}$ and the elemental matrix $\mathbf{H}$ are known, it suffices to use the combinatorial with the differential method to obtain the correct stoichiometric matrix, because the list of the feasible matrix is small and therefore there are a reduced number of combinations.

Nevertheless, in spite of having the correct stoichiometric matrix, the differential method does not give good values for the parameters in $\mathbf{k}$, since we are making an error assuming that the derivatives of the fitting curves are the fitting curves of the derivatives.

Plot of the concentrations obtained using the stoichiometric matrix (6.1), and $\mathbf{k}_{Alg1}$, and $\mathbf{k}_{Alg2}$ in Table 6.2 are given in Figures 6.3 and 6.4 respectively.

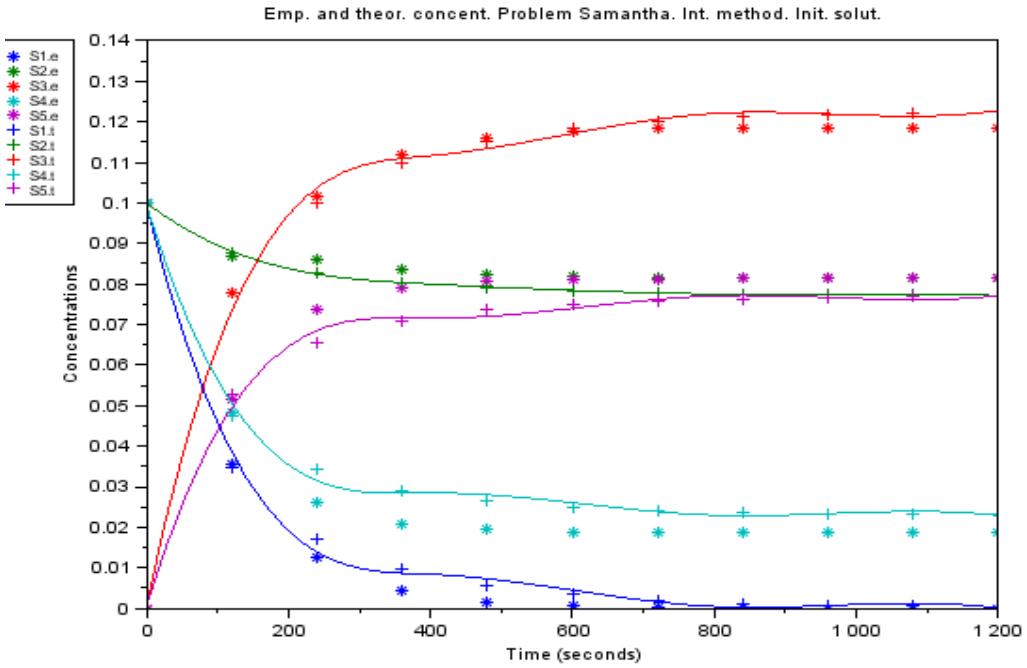

Figure 6.2: Empirical and theoretical concentrations of the data in [6] with the initial parameters and the differential method.

Stoichiometric matrix	$\mathbf{A}_{comb-diff} = \begin{pmatrix} 0 & 0 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & 2 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$
Alg 1. rate vector	$\mathbf{k}_{Alg1} = \begin{pmatrix} 0.0329696098 \\ 0.1164830894 \\ 0.1130988745 \end{pmatrix}$
Alg 1. obj value	$1.3154385597 \cdot 10^{-13}$
Alg 2. rate vector	$\mathbf{k}_{Alg2} = \begin{pmatrix} 0.0500029707 \\ 0.1499983701 \\ 0.1000005432 \end{pmatrix}$
Alg 2. obj value	$9.2939083610 \cdot 10^{-17}$

Table 6.2: Parameters for the Algorithms 1 and 2 for the data set A.1.

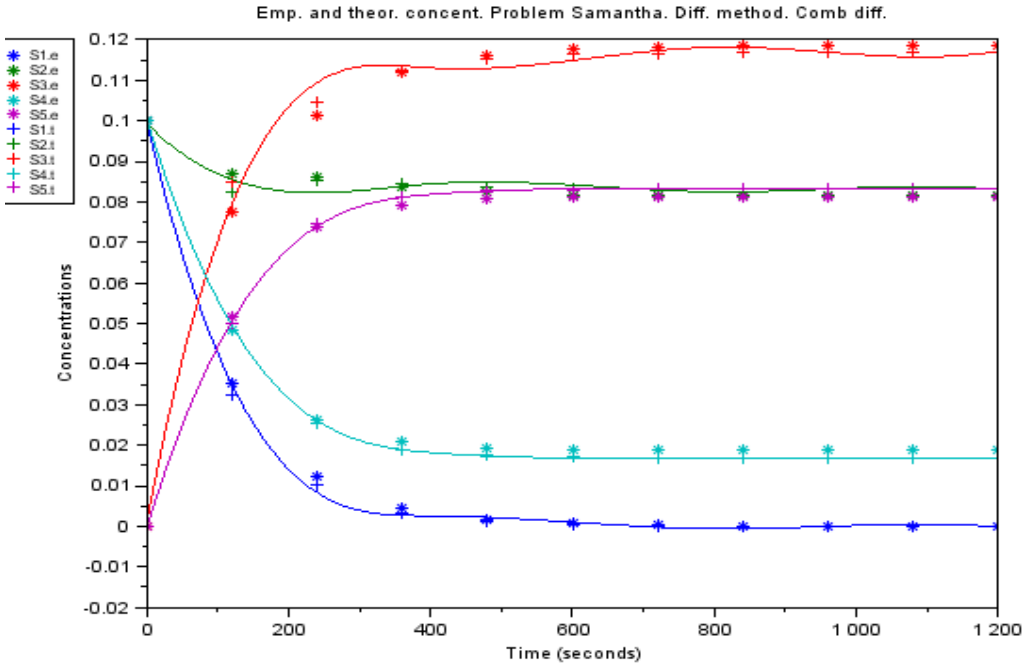

Figure 6.3: Empirical and theoretical concentrations of the data in [6] after the combinatory using the differential method and using the differential method

In Table 6.3 the different improvements in terms of objective value along the process can be seen. The optimal matrix has been soon trapped, that is the reason why the improvements cannot be well observed in Figure 6.8.

Observe that throughout this chapter, all tables with times start at 0 so they reflect the time involved in executing just the phase which is considered.

The last step in the strategy proposed (combinatory with integral method) is used to improve even more the results. In our case, that is impossible because the correct stoichiometric matrix is obtained. However, we cannot forget that overfitting may happen: there may exists a set of parameters that gives a better accuracy to the empirical data and they are not the ones we are looking for. As we can see in Table 6.4 that is not our case.

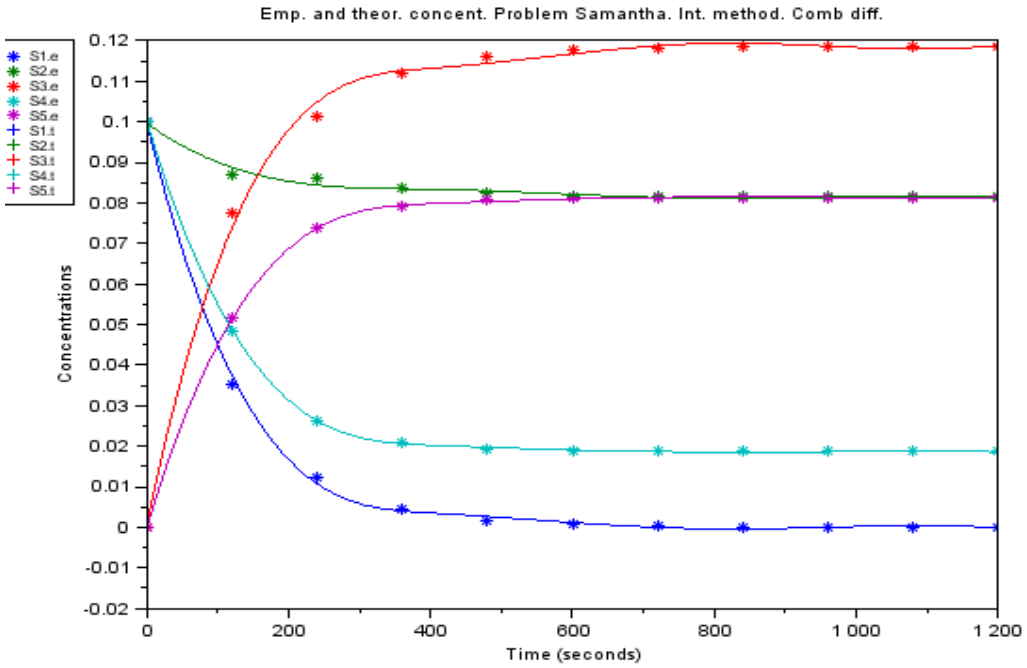

Figure 6.4: Empirical and theoretical concentrations of the data in [6] after the combinatory using the differential method and using the integral method

The concentrations using the parameters for the algorithms 3 and 4 (Table 6.4) are plotted in Figures 6.6 and 6.7 respectively.

Table 6.5 shows the improvements of the objective values along this last step. Graphically they can be seen in Figure 6.8.

6.2 Data Set 2

Let us begin this section by writing the matrix we take as initial solution in the data set A.2 and the corresponding rate reaction vector for the differential and the integral method (Table 6.6). Observe that in this case none of the correct reactions are chosen.

TIME (SECONDS)	OBJ. VALUE
0	$0.2049340758 \cdot 10^{-11}$
0.4380000000	$0.2049339312 \cdot 10^{-11}$
3.563000000	$0.2034930982 \cdot 10^{-11}$
5.141000000	$0.1358003582 \cdot 10^{-11}$
8.188000000	$0.1358003582 \cdot 10^{-11}$
8.281000000	$0.1315438572 \cdot 10^{-12}$
12.50000000	$0.1315438560 \cdot 10^{-12}$
17.98500000	$0.1315438560 \cdot 10^{-12}$
27.37500000	$0.1315438560 \cdot 10^{-12}$
27.89100000	$0.1315438560 \cdot 10^{-12}$
32.07800000	$0.1315438560 \cdot 10^{-12}$
92.65600000	$0.1315438560 \cdot 10^{-12}$
245.5160000	$0.1315438560 \cdot 10^{-12}$
1488.250000	$0.1315438560 \cdot 10^{-12}$
4059.469000	$0.1315438560 \cdot 10^{-12}$

Table 6.3: Improvements objective value in the combinatorial with the differential method for the data in [6].

Stoichiometric matrix	$\mathbf{A}_{comb-int} = \begin{pmatrix} 0 & 0 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & 2 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$
Algorithm 3 - rate vector	$k_{Alg3} = \begin{pmatrix} 0.1164837408 \\ 0.0329700477 \\ 0.1130988703 \end{pmatrix}$
Alg 3. obj value	$1.3154385597 \cdot 10^{-13}$
Algorithm 4 - rate vector	$k_{Alg4} = \begin{pmatrix} 0.1499983852 \\ 0.0500029710 \\ 0.1000005385 \end{pmatrix}$
Alg 4. obj value	$9.2939083610 \cdot 10^{-17}$

Table 6.4: Parameters for the Algorithms 3 and 4.

Plots of the theoretical concentrations obtained with the parameters of the differential and integral method are shown in Figures 6.9 and 6.10 respectively.

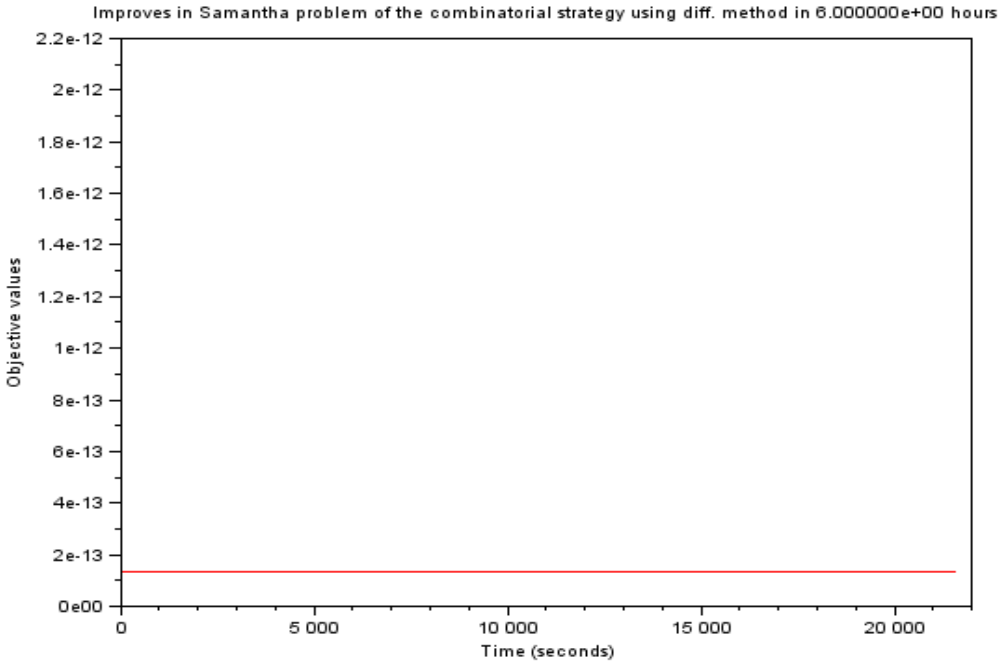

Figure 6.5: Improvements in the combinatorial with the differential method for the data in [6].

TIME (SECONDS)	OBJ. VALUE
0	$0.9293908361 \cdot 10^{-16}$
113.6410000	$0.9289020382 \cdot 10^{-16}$
4256.125000	$0.9288976172 \cdot 10^{-16}$
4331.328000	$0.9288964508 \cdot 10^{-16}$

Table 6.5: Improvements objective value in the combinatory with the integral method for the data in [6]

Data set A.2 differs from the previous example in the sense that vector $\mathcal{M}$ and matrix $\mathbf{H}$ are not given, so the set of feasible matrices is of much larger cardinality, and thus we have more combinations of reactions to test.

After the combinatory with the differential method the stoichiometric matrix obtained and the rate reaction vectors, $\mathbf{k}$, for the Algorithms 1 and 2 are in Table 6.7.

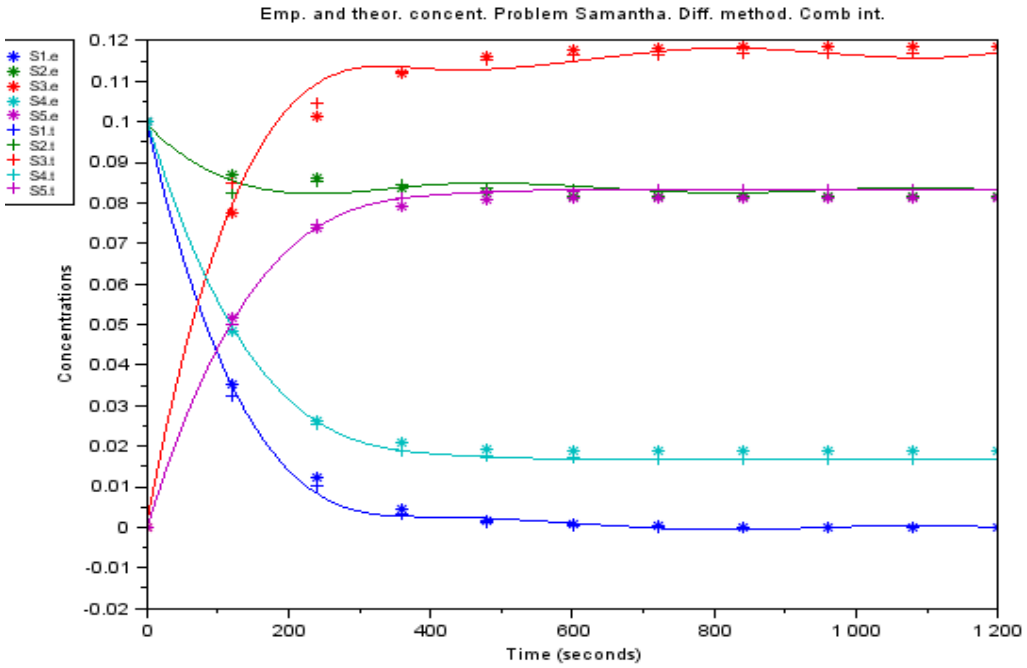

Figure 6.6: Empirical and theoretical concentrations of the data in [6] after the combinatory using the integral method and with the differential method

Plots of the concentration after the combinatory with differential method are shown in Figures 6.11 and 6.12.

Note that, in spite of not having the correct parameters, the fitting is almost perfect in some species.

The improvements along the process are shown in Table 6.8 and the plot in Figure 6.16.

Finally, let us see in Table 6.9 the matrix obtained in the combinatory with the integral method and the rate reaction vector calculated with this matrix and the algorithms 3 and 4.

Observe that the stoichiometric matrix obtained after the combinatorial with the integral method is exactly the same as the one obtained with the differential one given in Table 6.7. This is due to the fact, among others, of the absence of the molecular weights vector $\mathcal{M}$ and the elemental matrix H.

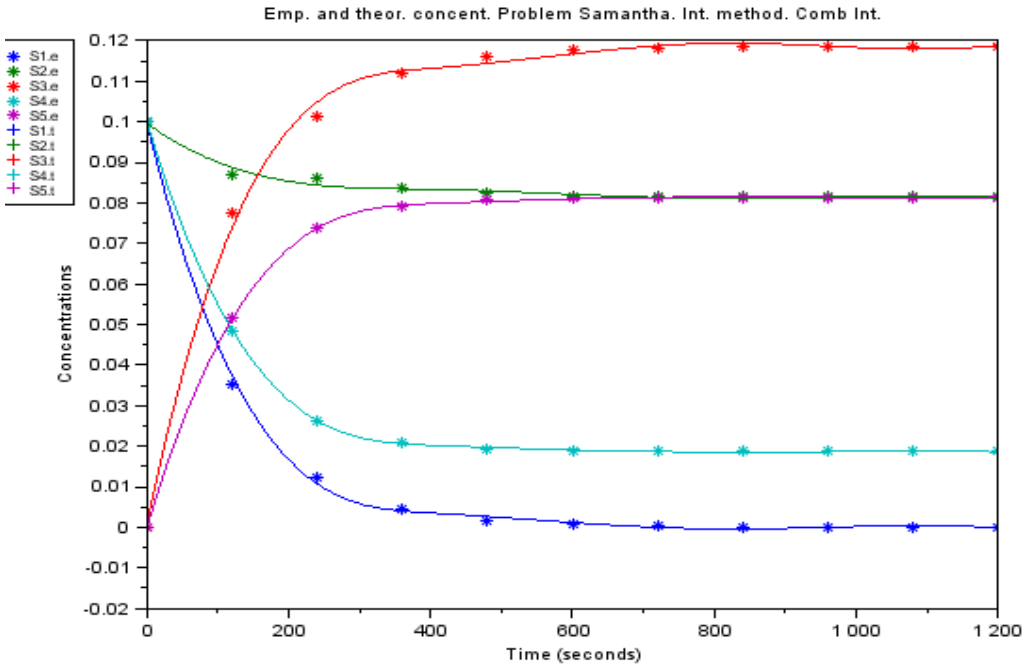

Figure 6.7: Empirical and theoretical concentrations of the data in [6] after the combinatory using the integral method and with the integral method

.

Concentrations plots of the results of Algorithms 3 and 4 can be seen in Figures 6.14 and 6.15. They will be very similiar to Figures (6.11) and (6.12).

Improvements along the process can be seen in Table 6.10 and Figure 6.16.

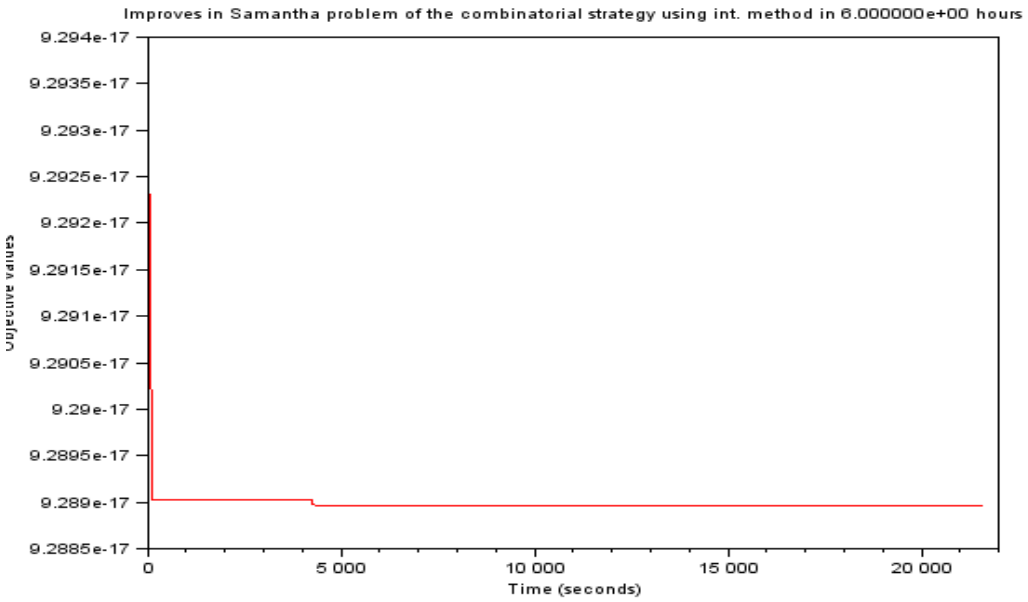

Figure 6.8: Improvements in the combinatory with the integral method for the data in [6]

.

Initial stoichiometric matrix	$\mathbf{A}_{ini} = \begin{pmatrix} -1 & 1 & -1 & -2 \\ 0 & 0 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$
Int. method rate vector	$\mathbf{k}_{ini-diff} = \begin{pmatrix} 0.0033319342 \\ 0.0000513809 \\ 0.0041930041 \\ 0.0032959845 \end{pmatrix}$
Diff. obj value	$5.6519473266 \cdot 10^{-10}$
Diff. method rate vector	$\mathbf{k}_{ini-int} = \begin{pmatrix} 0.0000100000 \\ 0.0169919665 \\ 0.0098788095 \\ 0.0010711874 \end{pmatrix}$
Int. obj value	$1.5335471601 \cdot 10^{-5}$

Table 6.6: Initial stoichiometric matrix and reaction vector for the differential and integral method in data set A.2.

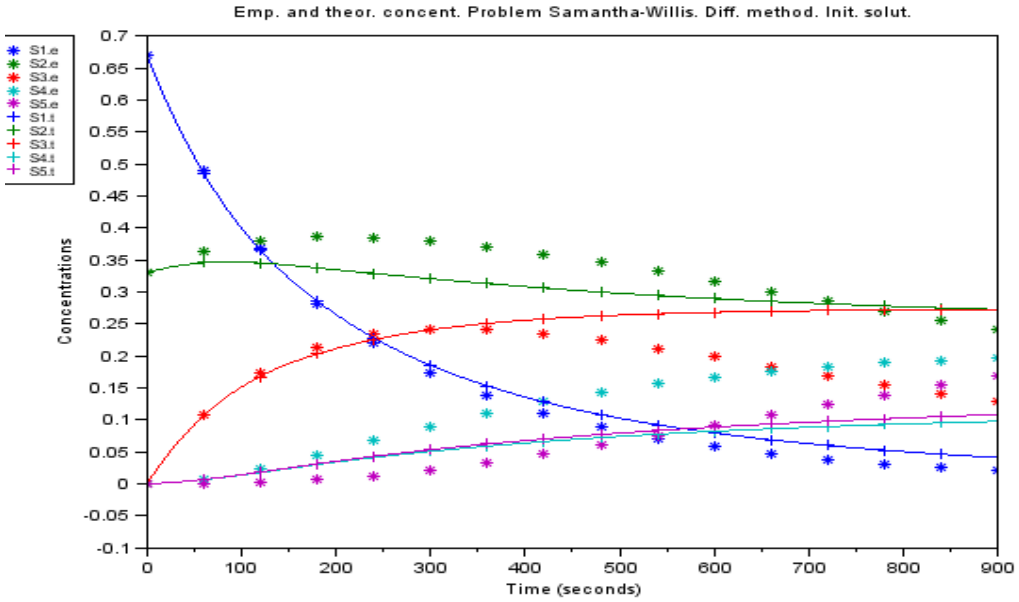

Figure 6.9: Empirical and theoretical concentrations of the data in [5] with the initial parameters and the differential method

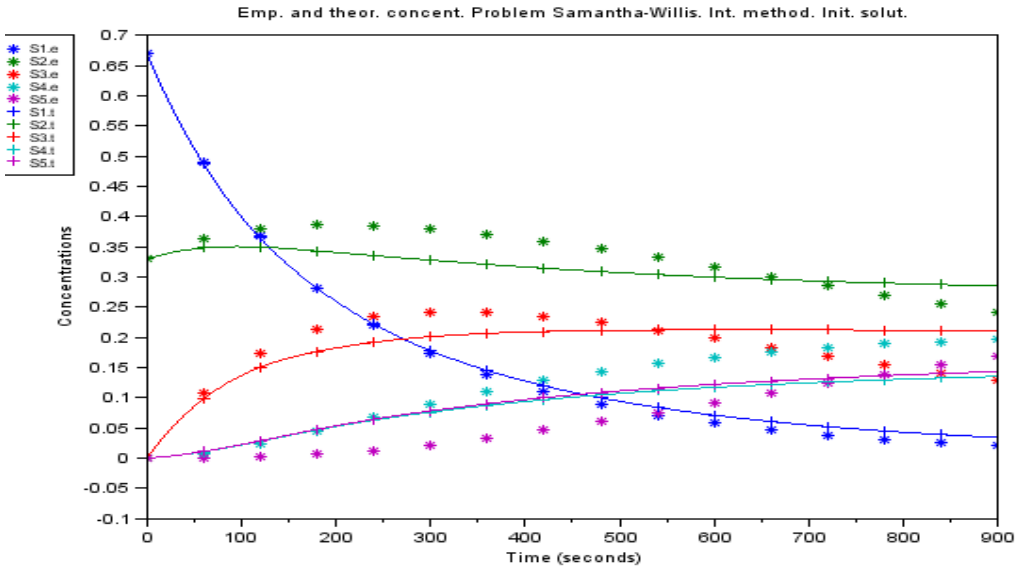

Figure 6.10: Empirical and theoretical concentrations of the data in [5] with the initial parameters and the integral method.

Stoichiometric matrix	$\mathbf{A}_{comb-diff} = \begin{pmatrix} 0 & -2 & -1 & -2 \\ -1 & 0 & -1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{pmatrix}$
Alg 1. rate vector	$\mathbf{k}_{Alg1} = \begin{pmatrix} 0.0024306747 \\ 0.2199761353 \\ 0.0030956085 \\ 0.0039886964 \end{pmatrix}$
Alg 1. obj value	$1.4693823086 \cdot 10^{-10}$
Alg 2. rate vector	$\mathbf{k}_{Alg2} = \begin{pmatrix} 0.0022631146 \\ 0.1672035153 \\ 0.0023254067 \\ 0.0037494357 \end{pmatrix}$
Alg 2. obj value	$4.2592913072 \cdot 10^{-6}$

Table 6.7: Parameters for the Algorithms 1 and 2 of the Data Set A.2.

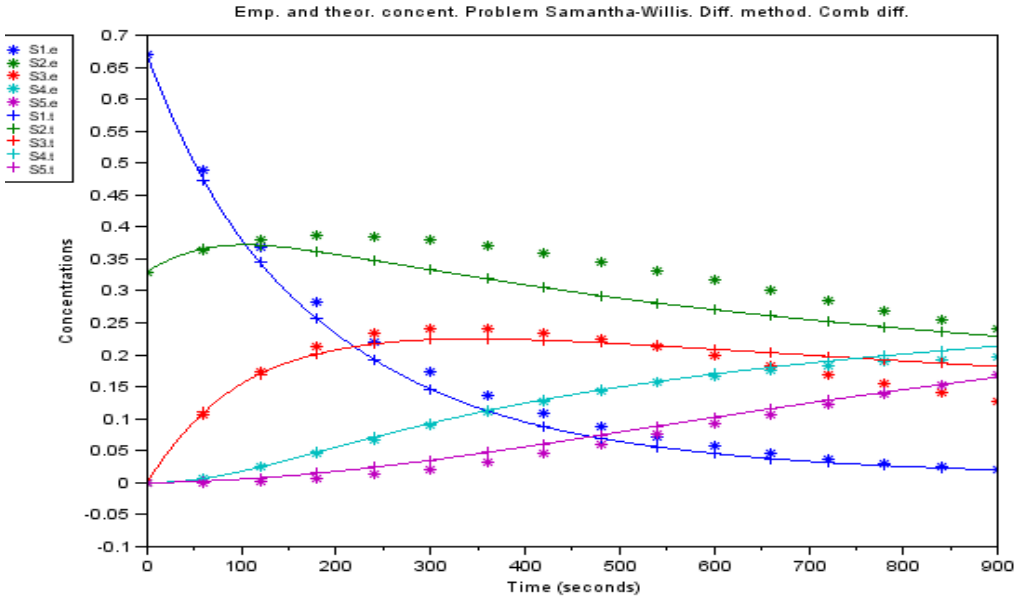

Figure 6.11: Empirical and theoretical concentrations of the data in [5] after the Algorithm 1.

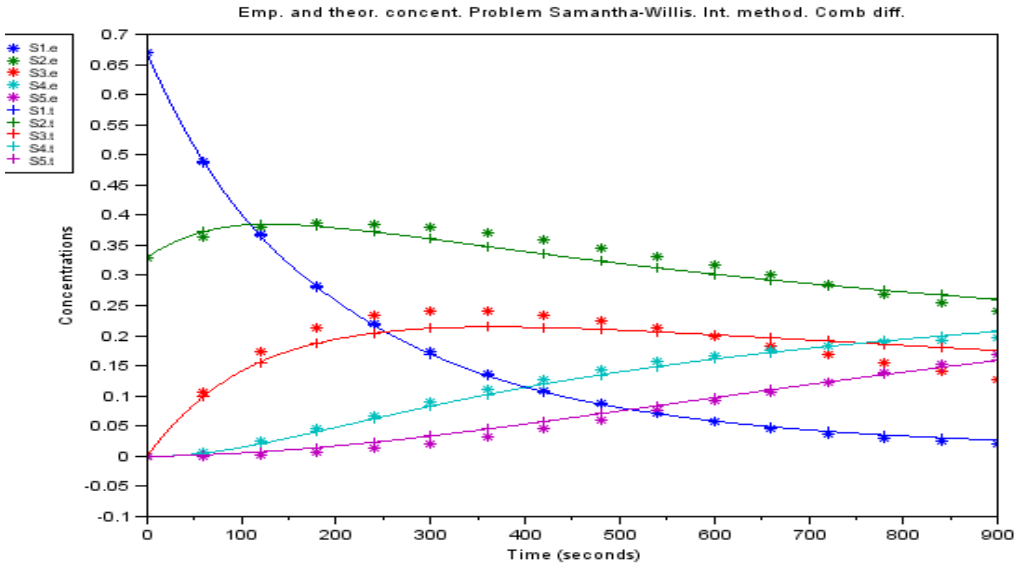

Figure 6.12: Empirical and theoretical concentrations of the data in [5] after the Algorithm 2.

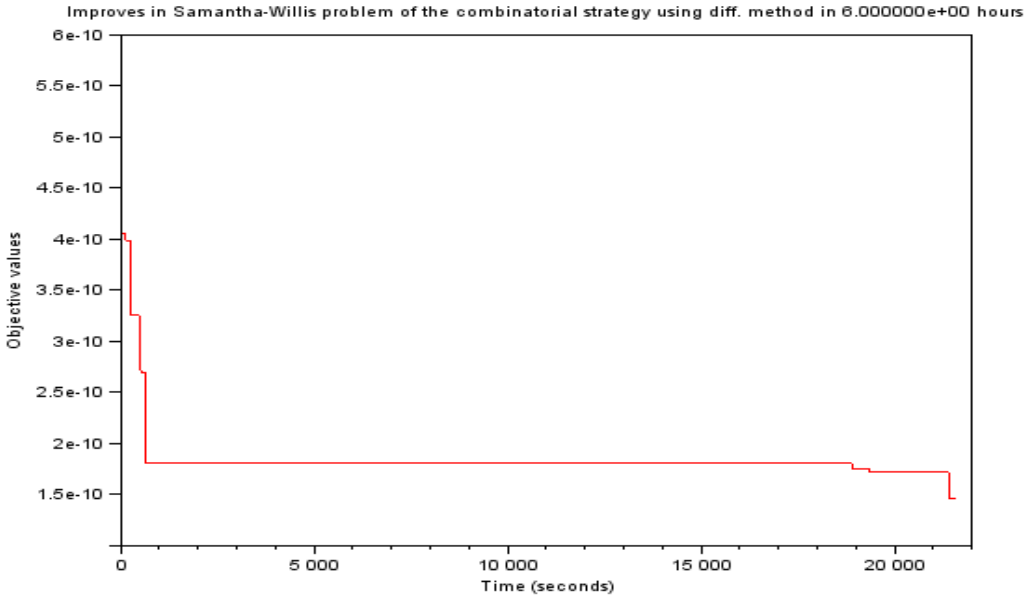

Figure 6.13: Improvements in the combinatory with the differential method for the data in [5].

TIME (SECONDS)	OBJ. VALUE
0	$0.5651947327 \cdot 10^{-9}$
3.062000000	$0.4635914666 \cdot 10^{-9}$
20.20300000	$0.4055557529 \cdot 10^{-9}$
129.9210000	$0.3991671705 \cdot 10^{-9}$
240.8430000	$0.3923546641 \cdot 10^{-9}$
249.8430000	$0.3254454054 \cdot 10^{-9}$
477.1090000	$0.3134104704 \cdot 10^{-9}$
479.2650000	$0.2850598601 \cdot 10^{-9}$
505.6090000	$0.2719036922 \cdot 10^{-9}$
526.5000000	$0.2692205611 \cdot 10^{-9}$
657.2960000	$0.2417530342 \cdot 10^{-9}$
660.4840000	$0.1804711967 \cdot 10^{-9}$
1433.921000	$0.1804711966 \cdot 10^{-9}$
3096.171000	$0.1804711966 \cdot 10^{-9}$
18903.46800	$0.1758743592 \cdot 10^{-9}$
19334.84300	$0.1725465483 \cdot 10^{-9}$
20681.45300	$0.1719223189 \cdot 10^{-9}$
21087.32800	$0.1719223189 \cdot 10^{-9}$
21400.87500	$0.1469382309 \cdot 10^{-9}$

Table 6.8: Improvements objective value in the combinatory with the differential method for the data in [5].

Stoichiometric matrix	$\mathbf{A}_{comb-int} = \begin{pmatrix} 0 & -2 & -1 & -2 \\ -1 & 0 & -1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{pmatrix}$
Alg 3. rate vector	$\mathbf{k}_{Alg3} = \begin{pmatrix} 0.0024306746 \\ 0.2199761474 \\ 0.0030956084 \\ 0.0039886964 \end{pmatrix}$
Alg 3. obj value	$1.4693823086 \cdot 10^{-10}$
Alg 4. rate vector	$\mathbf{k}_{Alg2} = \begin{pmatrix} 0.0021676152 \\ 0.2869917220 \\ 0.0023575863 \\ 0.0036784425 \end{pmatrix}$
Alg 4. obj value	$5.3409901223 \cdot 10^{-6}$

Table 6.9: Parameters for the Algorithms 3 and 4 of the Data Set A.2.

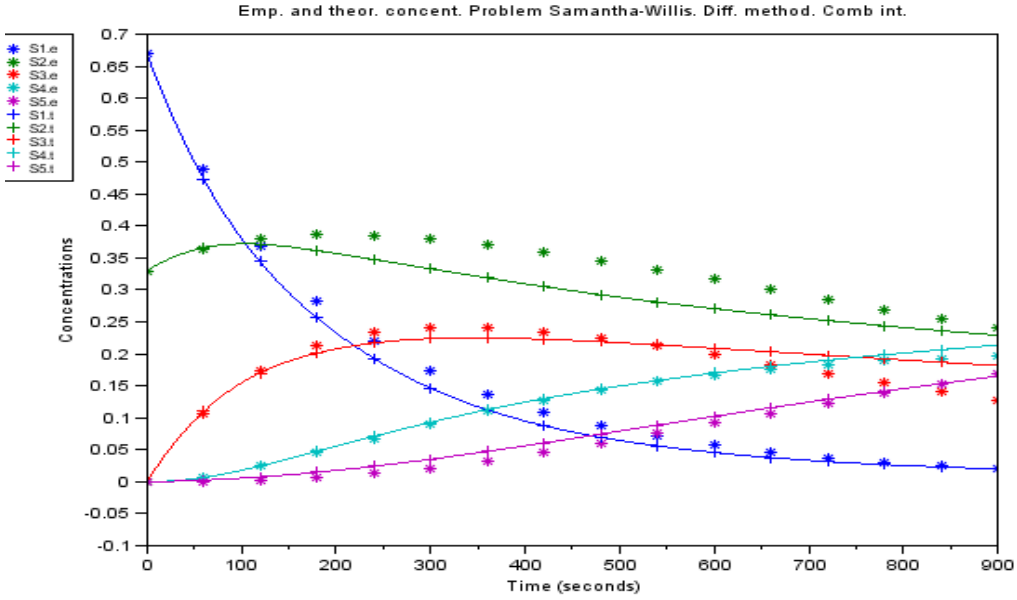

Figure 6.14: Empirical and theoretical concentrations of the data in [5] after the Algorithm 3.

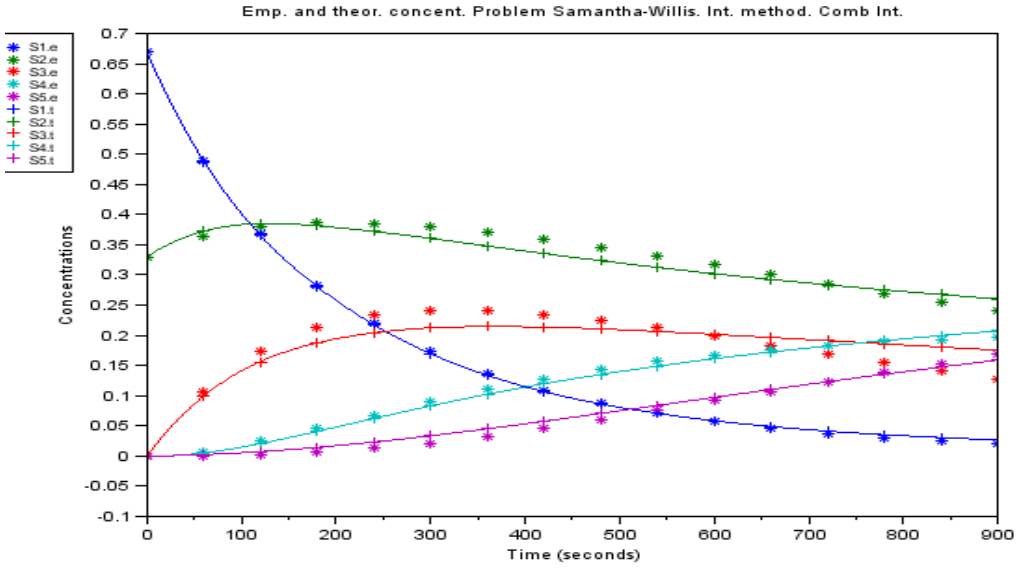

Figure 6.15: Empirical and theoretical concentrations of the data in [5] after the Algorithm 4.

TIME (SECONDS)	OBJ. VALUE
0	$0.4259291307 \cdot 10^{-5}$

Table 6.10: Improvements objective value in the combinatory with the integral method for the data in [5].

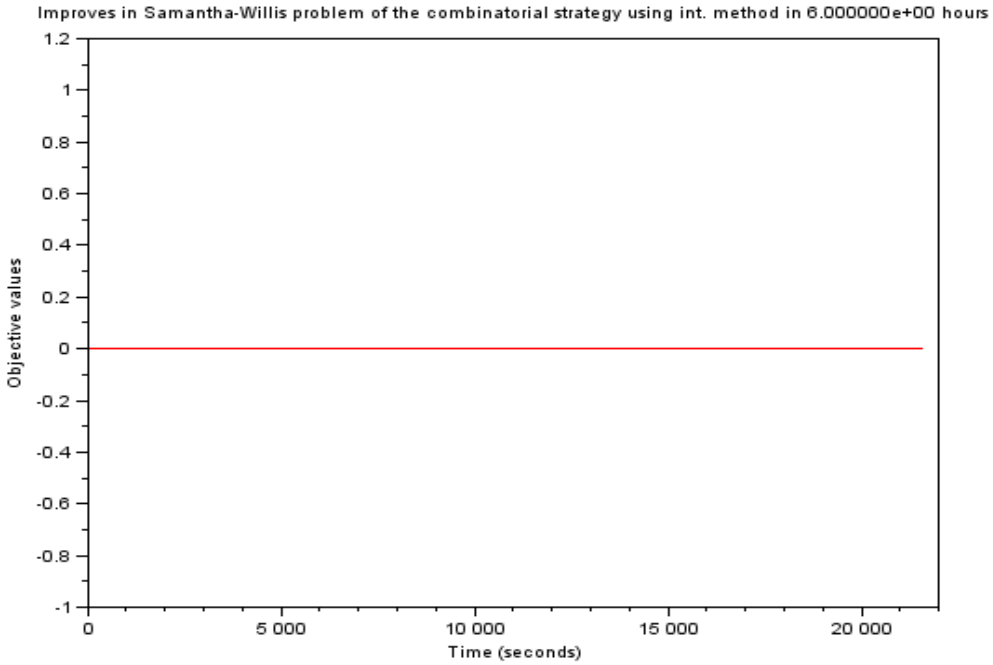

Figure 6.16: Improvements in the combinatory with the integral method for the data in [5].

Appendix A

Data Sets

A.1 Data Set 1

The first set of data is taken from [6]. We have:

- $N = 5$.
- $L = 3$.
- $N_0 = 11$.
- Time steps:

$$T = T_{1\times 11} = \begin{pmatrix} 0 & 120 & 240 & 360 & 480 & 600 & 720 & 840 & 960 & 1080 & 1200 \end{pmatrix}$$

- Elemental matrix:

$$\mathbf{H} = \mathbf{H}_{3\times 5} = \begin{pmatrix} 4 & 0 & 2 & 1 & 3 \\ 3 & 1 & 2 & 1 & 2 \\ 6 & 2 & 4 & 4 & 6 \end{pmatrix}$$

- Molecular weights vector:

$$\mathcal{M} = \mathcal{M}_{1\times 5} = \begin{pmatrix} 102 & 18 & 60 & 32 & 74 \end{pmatrix}$$

- Initial concentrations:

$$\mathbf{Y_0} = \mathbf{Y_0}_{1\times 5} = \begin{pmatrix} 0.1 & 0.1 & 0 & 0.1 & 0 \end{pmatrix}$$

SPECIES / TIME	1	2	3	4	5
0	0.1	0.1	0	0.1	0
120	0.0354075646011	0.0869966364218	0.0775957989771	0.0484109281793	0.0515890718207
240	0.0124453858664	0.0861603776298	0.1013942365038	0.0262850082366	0.0737149917634
360	0.0045008700064	0.0835515871445	0.1119475428491	0.0209492828619	0.0790507171381
480	0.0016663254963	0.0822421381718	0.1160915363319	0.0194241873245	0.0805758126755
600	0.0006233680361	0.0817130985407	0.1176635334233	0.0189102694954	0.0810897305046
720	0.0002341574053	0.0815097416442	0.1182561009505	0.0187244157611	0.0812755842389
840	0.0000880947678	0.0814326389772	0.1184792662551	0.0186554557906	0.0813445442094
960	0.0000331626438	0.0814035333430	0.1185633040132	0.0186296293007	0.0813703706993
1080	0.0000124866243	0.0813925630206	0.1185949503550	0.0186199236037	0.0813800763963
1200	0.0000047019440	0.0813884304655	0.1186068675905	0.0186162714784	0.0813837285216

Table A.1: Empirical concentrations of the data set in paper [6].

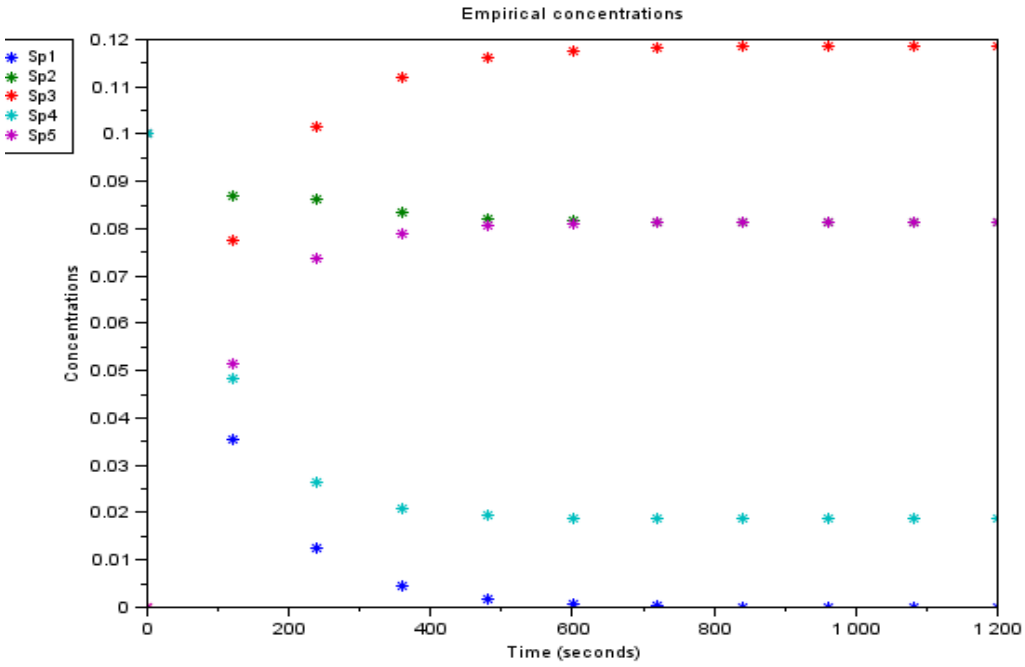

Figure A.1: Empirical Concentrations of the data set in [6].

The empirical concentrations of these data are shown in Table A.1. A plot of these concentrations can be seen in Figure A.1.

In this particular case the solution of the parameters **A** and **k** are known, as it can be seen below.

- Stoichiometric matrix:

$$\mathbf{A} = \mathbf{A}_{5\times 3} = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & -1 \\ 2 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

- Stoichiometric reactive matrix:

$$\mathbf{R} = \mathbf{R}_{5\times 3} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Stoichiometric product matrix:

$$\mathbf{P} = \mathbf{P}_{5\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Rate coefficients:

$$\mathbf{k} = \mathbf{k}_{3\times 1} = \begin{pmatrix} 0.1 \\ 0.15 \\ 0.05 \end{pmatrix}$$

A.2 Data Set 2

The second data set corresponds to the article [5].

- $N = 5$.
- $L = 4$.
- $N_0 = 16$.
- Time steps:

$$T = T_{1\times 16} = \begin{pmatrix} 0 & 60 & 120 & 180 & 240 & 300 & 360 & 420 & 480 & 540 & 600 & 660 & 720 & 780 & 840 & 900 \end{pmatrix}$$

- Initial concentrations:

$$\mathbf{Y_0} = \mathbf{Y_{0_{1\times 5}}} = \begin{pmatrix} 0.67 & 0.33 & 0 & 0 & 0 \end{pmatrix}$$

Empirical concentrations are shown in table A.2. Graphically, they can be seen in Figure A.2.

SPECIES / TIME	1	2	3	4	5
0	0.67	0.33	0	0	0
60	0.4891429	0.3628479	0.1071183	0.0072335	0.0002698
120	0.3678596	0.3793172	0.173477	0.0241317	0.0019658
180	0.2823506	0.3858145	0.2126880	0.0454913	0.0059471
240	0.2199138	0.3854987	0.2334394	0.0680772	0.0125241
300	0.1731481	0.3802419	0.2415418	0.0899689	0.0216191
360	0.1374479	0.3713392	0.2410032	0.1101102	0.0329201
420	0.1097974	0.3597800	0.2346443	0.1279989	0.0459998
480	0.0881404	0.3463562	0.2244757	0.143476	0.0603985
540	0.0710285	0.3317089	0.2119397	0.1565849	0.0756763
600	0.0574141	0.3163537	0.1980716	0.16748	0.0914423
660	0.0465226	0.3006980	0.1836121	0.1763684	0.1073669
720	0.0377710	0.2850577	0.1690856	0.1834753	0.1231843
780	0.0307141	0.2696705	0.1548571	0.1890232	0.1386882
840	0.0250074	0.2547105	0.1411737	0.1932212	0.1537256
900	0.0203819	0.2402994	0.1281955	0.1962598	0.1681880

Table A.2: Empirical concentrations of the data set in paper [5].

In contrast to the data set in Section A.1, matrices $\mathbf{H}$ and $\mathcal{M}$ are unknown. Nevertheless, the optimal stoichiometric matrix $\mathbf{A}$ and rate coefficient vector $\mathbf{k}$ are known as happened before.

- Stoichiometric matrix:

$$\mathbf{A} = \mathbf{A}_{5\times 4} = \begin{pmatrix} -2 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Stoichiometric reactive matrix:

$$\mathbf{R} = \mathbf{R}_{5\times 4} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

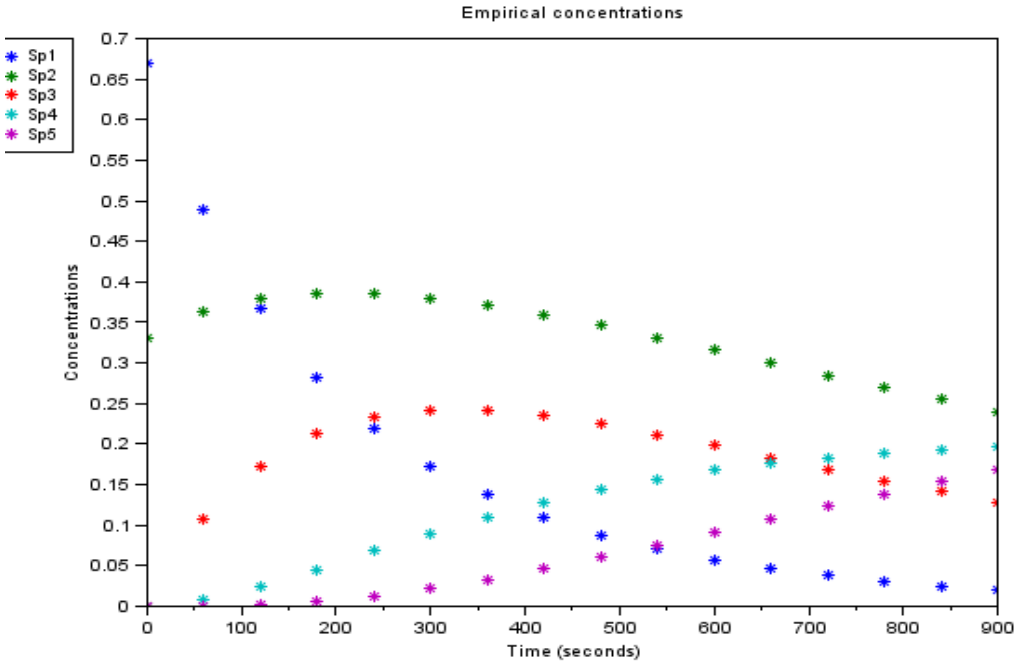

Figure A.2: Empirical Concentrations of the data set in [5].

- Stoichiometric product matrix:

$$\mathbf{P} = \mathbf{P}_{5\times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Rate coefficients:

$$\mathbf{k} = \mathbf{k}_{4\times 1} = \begin{pmatrix} 0.0016667 \\ 0.0033333 \\ 0.0021667 \\ 0.005 \end{pmatrix}$$

Bibliography

[1] André Bardow and Wolfgang Marquardt. Incremental and simultaneous identification of reaction kinetics: methods and comparison. *Chemical Engineering Science*, 59(1):2673–2684, 2004.

[2] Anna Bassi and Stefano De Marchi. A scilab radial basis functions toolbox.

[3] S Bermúdez, R Blanquero, JA Hernández, and J Planelles. A new parametric model for fitting fertility curves. *Population studies*, 66(3):297–310, 2012.

[4] Samuel Burer and Adam N Letchford. Non-convex mixed-integer nonlinear programming: A survey. *Surveys in Operations Research and Management Science*, 17(2):97–106, 2012.

[5] Samantha C. Burnham, Dominic P. Searson, Mark J. Willis, and Allen R. Wright. Inference of chemical reaction networks. *Chemical Engineering Science*, 63(1):862–873, 2008.

[6] Samantha C. Burnham and Mark J. Willis. Determining reaction networks. *Computer Aided Chemical Engineering*, 27(1):561–566, 2009.

[7] Scilab Consortium et al. Scilab manual, 2010.

[8] IBM ILOG Cplex. 12.2 users manual, 2010.

[9] Carl De Boor. *A practical guide to splines*, volume 27. Springer-Verlag New York, 1978.

[10] Sagar B Gadewar, Michael F Doherty, and Michael F Malone. A systematic method for reaction invariants and mole balances for complex chemistries. *Computers & Chemical Engineering*, 25(9):1199–1217, 2001.

[11] Peter Gans. *Data fitting in the chemical sciences: by the method of least squares*. John Wiley & Sons Inc, 1992.

[12] Ralf Hanneman-Tams, Attila Gbor, Gbor Szederknyi, and Katalin M. Hangos. Model complexity reduction of chemical reaction networks using mixed-integer quadratic programming. *Computers and Mathematics with Applications*, 65(1):1575–1595, 2012.

[13] Pierre Hansen and Nenad Mladenović. Variable neighborhood search: Principles and applications. *European Journal of Operational Research*, 130(3):449–467, 2001.

[14] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, pages 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.

[15] Nenad Mladenović and Pierre Hansen. Variable neighborhood search. *Computers & Operations Research*, 24(11):1097–1100, 1997.

[16] BA Murtagh and MA Saunders. Minos user's manual, 1985.

[17] William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[18] Dominic P. Searson, Mark J. Willis, and Allen Wright. Reverse engineering chemical reaction networks from time series data. *Statistical Modelling of molecular descriptors in QSAR/QSPR*, 7(1):327–348, 2012.

[19] Dominic P. Searson, Mark J. Willis, Allen R. Wright, and Simon J. Horne. Inference of chemical reaction networks using hybrid s-system models. *Chemical Product and Process Modeling*, 2(1):862–873, 2007.

[20] JM Varah. A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific and Statistical Computing*, 3(1):28–46, 1982.