10th International IFAC Symposium on
Computer Applications in Biotechnology
Preprints Vol.1, June 4-6, 2007, Cancún, Mexico

USE OF MODULATING FUNCTIONS FOR REACTION NETWORK IDENTIFICATION

Olivier Bernard

INRIA-COMORE, BP93,
06902 Sophia-Antipolis Cedex, France,
olivier.bernard@inria.fr

Abstract: This paper improves a methodology that determines the pseudo-stoichiometric coefficient matrix K in a mass balance based model. The idea consists in using modulating functions in order to filter the data that are often noisy for biotechnological processes. The first stage consists in estimating the number of reactions that must be taken into account to represent the main mass transfer within the bioreactor. This provides the dimension of K. Then a set of modulating functions are used to directly determine the matrix coefficients. This method is illustrated with simulations of a process of lipase production from olive oil by *Candida rugosa* and compared to the method of Bernard and Bastin (2005a) based on a temporal approach.

Keywords: Modelling, Identification, Nonlinear systems, Bioreactors

1. INTRODUCTION AND MOTIVATION

A key challenge emerging in biotechnology consists in identifying a model from a large set of data when the underlying process mechanisms are unclear. This is especially the case for data issued from large networks like metabolic or genetic networks. For example, hundreds of genes can dynamically interact in the cell to finally trigger or inhibit the production of a metabolite. Here we will focus on mass balance based systems that are characterized by mass fluxes (and not information fluxes) between several compartments. In this context the question of identifying a mass balance based model can arise mainly for two reasons. Classically, the first class of problems concerns complex systems with many interactions between species that may not all have been accurately identified. This is typically the case for a natural ecosystem where there are continuous species successions. It can also appear for some biotechnological complex processes. For example wastewater treatment processes often involve bacterial consortium made of a broad range of bacterial species degrading a mixture of organic substrates. For instance, more than 140 bacterial species have been found (Delbès et al. (2001)) in an anaerobic digestion wastewater treatment plant. Identifying a simplified macroscopic model that can describe these processes is thus an important issue. The second important motivation is radically different. It addresses the question of model reduction. When one wants *e.g.* to design online algorithms for bioreactor monitoring, control and optimisation (Bastin and Dochain (1990)), a model that can be mathematically handled may be required. Reducing a complicated given reaction network to a much simpler model also can be targeted by the proposed approach.

For these two purposes, our objective is to end-up with the following general macroscopic mass-balance model:

$$\frac{d\xi(t)}{dt} = K\, r(t) + v(t), \qquad (1)$$

In this model, the vector $\xi = (\xi_1, \xi_2, \ldots, \xi_n)^T$ is made-up of the concentrations of the various species inside the liquid medium. The term $v(t)$ represents the net balance between inflows, outflows and di-

lution effects. The term $K\,r(t)$ represents the biological and biochemical conversions in the reactor (per unit of time) according to some underlying reaction network. The $(n \times p)$ matrix K is a constant pseudo-stoichiometric matrix. $r(t) = (r_1(t), r_2(t), \ldots, r_p(t))^T$ is a vector of reaction rates (or conversion rates). It is supposed to depend on the state ξ and on external environmental factors.

The pseudo-stoichiometric (PS) matrix K is associated to a macroscopic reaction network that lumps together the many intracellular metabolic reactions of the various involved microbial species. The reaction network then summarises the main mass transfer throughout the bioreactor by a few reactions involving mainly extracellular compounds and biomasses without describing the intracellular behaviour. Each column of the matrix corresponds to a chemical or biological reaction of the underlying macroscopic reaction network. The coefficients $k_{ij} \;\; j = 1, \ldots, p$ are associated with the j^{th} reaction. A positive k_{ij} means that the i^{th} species ξ_i is a product of the j^{th} reaction, while a negative k_{ij} means that ξ_i is a substrate of the j^{th} reaction. If $k_{ij} = 0$ the species ξ_i is not involved in the j^{th} reaction.

In this paper, we are concerned with modelling situations where the on-line concentrations ξ_i of the involved species are measured but the structure of the simplified reaction network is *a priori* questionable and therefore the matrix K is partially unknown. The objective, as in (Bogaerts and Vande Wouwer (2001)), is to provide guidelines to the user for the identification of the structure of a macroscopic reaction network and the determination of the PS matrix K from a set of available data.

When the method is applied in order to simplify a known detailed intracellular metabolic network the concentrations ξ_i result from simulations of the detailed model that must be reduced.

The usual approach dedicated to the determination of reaction networks relies on the linearisation of the dynamics around a reference solution (Eiswirth et al. (1991); Chevalier et al. (1993)) and identification of the local Jacobian matrix. This approaches are then suitable for data close to steady state. Here, in the spirit of (Bernard and Bastin (2005b)), we use linear algebraic properties to exploit the structure of the bioprocesses (Equation (1)) and our arguments do not rely on any linearisation. As a consequence we are not limited to steady state data and we can exploit all the available measurements, even when associated to transient states.

Example: As in (Bernard and Bastin (2005a)), we consider the example of the production of lipase from olive oil by *Candida rugosa*. Here the microorganism is supposed to grow on two substrates that are produced by the hydrolysis of a primary complex organic substrate (olive oil), leading thus to the following 3-step reaction network (Chen and Bastin (1996)):

- Hydrolysis:

$$k_{11}S_1 \;+\; E \;\longrightarrow\; S_2 \;+\; k_{31}S_3 \;+ E$$

- Growth on S_2:

$$k_{22}S_2 \;+\; k_{62}O \;\longrightarrow\; X \;+\; k_{72}P$$

- Growth on S_3:

$$k_{33}S_3 \;+\; k_{63}O \;\longrightarrow\; X \;+\; k_{43}E \;+\; k_{73}P$$

where S_1 is the primary substrate (olive oil, made of several compounds, mainly triglycerides), S_2 (glycerol) and S_3 (fatty acids) are the secondary substrates. E is the enzyme (lipase), X the biomass (*Candida rugosa*), O the dissolved oxygen and P the dissolved carbon dioxide.

The associated PS matrix is:

$$K = \begin{pmatrix} -k_{11} & 0 & 0 \\ 1 & -k_{22} & 0 \\ k_{31} & 0 & -k_{33} \\ 0 & 0 & k_{43} \\ 0 & 1 & 1 \\ 0 & -k_{62} & -k_{63} \\ 0 & k_{72} & k_{73} \end{pmatrix} \quad \text{with } k_{ij} > 0.$$

We shall assume that this reaction network is unknown to the user and has to be discovered from data of the species concentrations. In order to validate the approach, the data will be simulated by a model but, of course, in practice the data are obtained from experiments.

Generally, the choice of a reaction network and its associated PS matrix K results from modelling assumptions. Sometimes however, a complete description of the reaction network is *a priori* not available. This can be a consequence of a lack of phenomenological knowledge on some of the involved mechanisms, rendering a part of the reaction network questionable. The problem can also arise when it is desired to reduce a complicated given reaction network to a much simpler model.

We first recall a method to determine the size of matrix K *i.e.* the number of independent reactions that are distinguishable from the available data. We explain how this methodology can be extended to better deal with noise filtering using modulating functions (MF). Then we show how the structure of matrix K can be estimated on the basis of the decomposition of the measurements on a specific basis. By structure we mean the sign and the location of the non-zero entries of matrix K. In addition, the method can also provide an estimate of the parameters k_{ij} if the available knowledge is sufficient. We compare this new approach with the one presented in (Bernard and Bastin (2005a)) based on a temporal approach.

2. EXPERIMENTAL DETERMINATION OF THE PS MATRIX RANK

2.1 Motivation

In this section, we recall how to determine the minimum number of reactions which are needed in order to explain the measured process behaviour. We assume that the vectors $\xi(t)$ of species concentrations and $v(t)$ of inflow/outflow balances are measured during some time interval $[t_0\ t_f]$ and exhibit significant variations with time. We assume also that the number of measured variables is larger than the number of reactions: $n > p$. The PS matrix K and the vector of reaction/conversion rates $r(t)$ are unknown. Note that we assume K to be a full rank matrix. Otherwise, it would mean that the same dynamical behaviour could be obtained with a matrix K of lower dimension, by defining other appropriate reaction rates.

2.2 Theoretical determination of $\dim(\mathcal{I}m(K))$ with a temporal approach (recall)

The model equation (1) can be viewed as a linear dynamical system with state ξ and inputs $r(t)$ and $v(t)$ (although we know obviously that r and v are state dependent).

$$\frac{d\xi(t)}{dt} - v(t) = Kr(t) \tag{2}$$

which has the form:

$$u(t) = Kw(t) \tag{3}$$

with $u(t) = \frac{d\xi(t)}{dt} - v(t)$ and $w(t) = r(t)$. However, the computation of u relies on the estimate of a derivative and is thus sensitive to noise. To improve this scheme, a first approach (Bernard et al. (1999); Bernard and Bastin (2005b,a)), consists in applying a linear filter in order to clean the data (noise reduction, decrease of autocorrelations etc ...) and to still end up with an expression of the form (3). For example, the moving average is a very simple filter that can be applied to (1) and leads to (3)(Bernard and Bastin (2005a)).

In this temporal approach we will compute $u(t)$ at various successive time instants t_i.

The question of the dimension of matrix K is then formulated as the estimation of the dimension of the image of K. In other words, the dimension of the space where $u(t)$ lives. The determination of the dimension of the $u(t)$ space is a classical problem in statistical analysis. It corresponds to the principal component analysis (see e.g. Horn and Johnson (1993)) that determines the dimension of the vector space spanned by the vectors k_i which are the rows of K. To reach this objective, we consider the $n \times N$ matrix U obtained from a set of N estimates of $u(t)$ at the time instants t_i:

$$U = (u(t_1), \ldots, u(t_N))$$

We will also consider the associated matrix of reaction rates, which is unknown:

$$W = (w(t_1), \ldots, w(t_N))$$

We consider more time instants t_i than state variables: $N > n$.

Property 1. For a matrix K of rank p, if W has full rank, then the $n \times n$ matrix $M = UU^T = KWW^TK^T$ has rank p, it has thus p positive eigenvalues $\sigma_i > 0$, and $(n-p)$ zeros eigenvalues.

Moreover, the eigenvectors associated with the σ_i generate an orthonormal basis of $\mathcal{I}mK$.

This property is a direct application of the singular decomposition theorem (Horn and Johnson (1993)).

Now from a theoretical point of view, it is clear that the number of reactions can be determined by counting the number of non zero singular values of UU^T.

In practice there are no zero eigenvalues and a statistical approach must be used. In order to give the same weighting to all the variables, the data vectors $u(t_i)$ must be normalised (Bernard and Bastin (2005a)). Then the proportion of the eigenvalues σ_i with respect to $\sum_j \sigma_j$ can be interpreted in term of variance associated with the corresponding eigenvector (inertia axis) (Horn and Johnson (1993)).

Now the question is to determine the number of eigenvectors that must be taken into account in order to produce a reasonable approximation of the data $u(t)$. The method consists in selecting the p first principal axis which represent a total variance larger than a fixed confidence threshold. For instance, in the next example, we will consider a threshold (depending on the information available on noise measurements) at 95% of the variance. This leads to the selection of 3 axis, and therefore $p = 3$.

Remark: If $\mathrm{rank}\ (M) = n$, it means that $rank(K) \geq n$. In such a case we cannot estimate p and measurements of additional variables are requested in order to apply the method presented here.

2.3 Theoretical determination of $\dim(\mathcal{I}m(K))$ using modulating functions

Here we propose to test an alternative approach, using a set of modulating functions (Shinbrot (1957); Preisig and Rippin (1993)) ϕ_i ($i \in \{1...N\}$) that are $\mathcal{C}^1$ and such that:

$$\phi_i(t_0) = \phi_i(t_f) = 0 \text{ for } i \in \{1...N\}$$

An application $f(t) : [t_0\ t_f] \rightarrow \mathbf{R}^n$ is modulated by taking the inner product with a MF ϕ

$$< f, \phi >= \int_{t_0}^{t_f} f(\tau)\phi(\tau)d\tau$$

Property 2. The following property holds:

$$< \frac{d\xi}{dt}, \phi >= - < \xi, \frac{d\phi}{dt} >$$

Proof: The proof is straightforwardly obtained after integrating by part the term $\int_{t_0}^{t_f} \frac{\xi}{dt}(\tau)\phi(\tau)d\tau$.

As a consequence, if we consider the inner product with any element of the set of MFs, we will have:

$$- < \xi(t), \frac{d\phi_i}{dt}(t) > - < v(t), \phi_i(t) >= K < r(t), \phi_i(t) >$$

Let us denote by $u_{\phi_i} = - < \xi, \frac{d\phi_i}{dt} > - < v, \phi_i >$ and $w_{\phi_i} =< r, \phi_i >$, we end up with an expression similar to equation (3):

$$u_{\phi_i} = K w_{\phi_i} \tag{4}$$

Now considering matrix U_ϕ made of the collection of vectors u_{ϕ_i} for the N considered MFs:

$$U_\phi = (u_{\phi_1}, \ldots, u_{\phi_N})$$

We will also consider the associated matrix of reaction rates, which is unknown:

$$W_\phi = (w_{\phi_1}, \ldots, w_{\phi_N})$$

These matrices verify $U_\phi = K W_\phi$. The same PCA approach as in paragraph 2.2 can thus be applied, leading thus to determination of both the number of nonzero eigenvalues and of the associated eigenvectors.

2.4 Example: the lipase production

We come back to the example which has been introduced above. Consistently with Equation (1), the model for the state

$$\xi = (S_1, S_2, S_3, E, X, O, P)^t$$

involving 3 main reactions can thus be written:

$$\frac{d\xi}{dt} = K \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} + v(t)$$

where $v(t) = D(\xi_{in} - \xi) - Q(\xi)$, with $\xi_{in} = (S_{1in}, S_{2in}, S_{3in}, 0, 0, 0, 0)^t$ the vector of influent concentrations, D is the dilution rate and $Q(\xi) = (0, 0, 0, 0, 0, qo_2(O), qco_2(P))^t$ the vector of gaseous flow rates.

Matrix K (that will be reconstructed from data) was chosen as follows:

$$K = \begin{pmatrix} -3 & 0 & 0 \\ 1 & -5 & 0 \\ 0.3 & 0 & -0.5 \\ 0 & 0 & 0.2 \\ 0 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0.3 & 1.5 \end{pmatrix}$$

For the simulation purpose, we assume that the kinetics of the three reactions are the same than in (Bernard and Bastin (2005a)).

A 30 day run of the model has been performed and the collected data have been corrupted with a multiplicative white noise of high magnitude (15% of the state) and resampled. Finally 667 data points are available (note that the considered noise is much higher than in Bernard and Bastin (2005a)).

An example of the collected data (after sampling) is presented in Figure 1. The state variables S_2, S_3, E, X, P, O and the gaseous flow rates q_{O_2} and q_{CO_2} have been measured. We assume here that the state variable S_1 was not recorded in order to illustrate the fact that our approach is applicable even if the full set of state variables is not available for measurement.

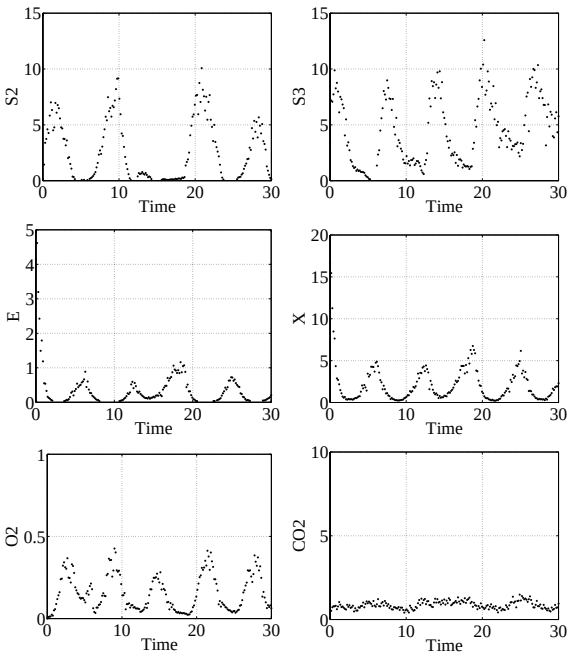

Fig. 1. Experiment simulated from the kinetic modelling corrupted with a multiplicative white noise.

The vectors u_{ϕ_i} are then computed using a Fourier basis of MF. For sake of comparison, the temporal $u(t_i)$ are also computed by applying a simple moving average. Finally, the eigenvectors of UU^T are computed.

Figure 2 represents the cumulated variance associated with the number of considered inertia axis. The results obtained with the temporal approach do perfectly correspond with what was expected: two reactions are sufficient to explain 82% of the observed variance and three reactions explain 95% of the total variance.

Surprisingly, the approach based on MFs has a very different result: one reaction seems to be able to explain 96% of the variance. This result is somewhat disappointing, but it can be easily explained. In the MF approach the vectors w are all multiplied by the same function $\phi(t)$. This induces thus a strong correlation between the components of w_ϕ. The result finally reproduces this correlation. Even if the MF approach is inefficient for assessing the reaction number, it will be

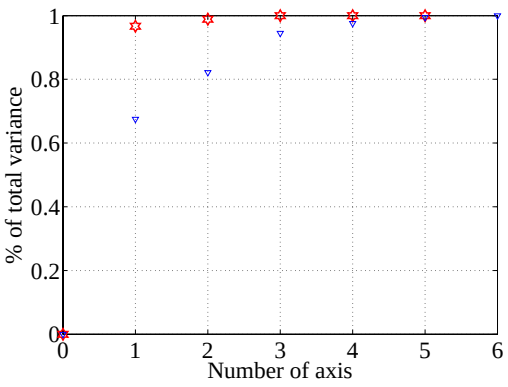

Fig. 2. Total variance explained with respect to the number of reactions for lipase production. Computation using a temporal analysis $\triangledown$ or MF $\star$.

seen in the sequel that it works much better than the temporal approach for the identification of matrix K.

3. ESTIMATION OF THE PSEUDO-STOICHIOMETRIC MATRIX K

We have previously estimated the number of involved reactions, we are thus in a position to start the estimation of the (totally or partially) unknown matrix K.

3.1 Determination of $\mathcal{I}mK$

From Property 1 we know that $\mathcal{I}mK$ is spanned by the eigenvectors ρ_i associated with the non zero eigenvalues of UU^T.

It means that each column k_i of K is a linear combination of the ρ_i. In other terms, there exists a $p \times p$ matrix G such that

$$K = \rho\, G$$

where the columns of matrix ρ are the eigenvectors ρ_j. In other words, the family of possible PS matrices K is parameterised by the $p \times p$ matrix G.

Remark 1: In general, since the reaction rates are unknown, matrix G (and therefore matrix K) is not identifiable (see Bernard and Bastin (2005a))), and one must add p^2 constraints to solve the problem.

Remark 2: The identification of each column K_k of K is independent of the identification of the other columns of K. indeed, let us call G_k the k^{th} column of G, we have $K_k = \rho G_k$. In the sequel we will thus focus on the identification of K_k, the kth column of K.

In order to make vector G_k (and thus K_k) uniquely identifiable, we need to introduce p additional structural constraints. At this stage, all the *a priori* knowledge on the reaction network should be considered to improve the estimation process. See (Bernard and Bastin (2005a)) for a review of the possible constraints that can be taken into account.

First, we impose that the reaction rate is normalised with respect to one species, and therefore that K_k contains one +1 or one -1. Note that sometimes we may not know the sign of the element: the two possible cases must then be considered.

It can be known for example that some components are not implied in a reaction, or that two yield coefficients must be equal (*e.g.* for the consumption of oxygen and the production of carbon dioxide). Finally, let us assume that p constraints are available for the k^{th} reaction. The following property explains how to compute the remaining coefficients of vector K_k.

Property 3. Let us assume that the p elements $k_{i_1,k}$ to $k_{i_p,k}$ of K_k are known. Vector K_k can then be computed from ρ, the basis of $\mathcal{I}m(K)$, as follows:

$$K_k = \rho\tilde{\rho}^{-1}\tilde{K}_k \tag{5}$$

where the subvector $\tilde{K}_k$ [resp. submatrix $\tilde{\rho}$] is extracted from K_k [resp. ρ] using the rows i_1 to i_p containing the known coefficients.

Proof: These submatrices obviously verify : $\tilde{K}_k = \tilde{\rho}G_k$. Since $\tilde{\rho}$ is a square matrix, it can generically be inverted, and thus provide vector $G_k = \tilde{\rho}^{-1}\tilde{K}_k$.

We emphasize that the result is independent of the other constraints on the other columns of K.

3.2 Example (continued)

In the considered example we consider that the first reaction is known, and we determine the two other reactions on the basis of experiments involving only S_2 and S_3, O, X, E and P. We are thus in the process of estimating the submatrix $\bar{K}$ extracted from K by removing the first row and the first column. The eigenvectors ρ_1 and ρ_2 associated with the two largest computed eigenvalues, following Property 1, are then the basis of $\mathcal{I}mK$.

To constrain the problem, we remark that a reaction still takes place when only S_2 [resp. S_3] is present at the initial time, and no S_3 [resp. S_2] is produced. In other words this means that S_2 is the only substrate of one reaction and that S_3 is the only substrate of the other one. Thus we will impose $\bar{k}_{12} = 0$ and $\bar{k}_{21} = 0$. Moreover we normalize the reaction with respect to biomass: $\bar{k}_{14} = \bar{k}_{41} = 1$. Let us focus on the reaction taking place in absence of S_3. The constraint submatrix $\tilde{\bar{K}}_1$ is obtained keeping only rows 2 and 4 of $\bar{K}_1$: $\tilde{\bar{K}}_1 = (0\ \ 1)^T$. The corresponding matrix $\tilde{\rho}$ can then be computed by extracting the second and fourth rows of ρ. Finally the first column of $\bar{K}$ can be computed: $\bar{K}_1 = \rho\tilde{\rho}^{-1}\tilde{\bar{K}}_1$.

The same methods leads straightforwardly to the computation of $\bar{K}_2$.

We shall now illustrate the proposed approach with a simulation study of the lipase production process. We assume that the first reaction is known, and therefore we only focus on the two other reactions.

$\bar{K}$	**temporal based identification** $\bar{K}$	**MF based identification** $\bar{K}$
$\begin{pmatrix} -5 & 0 \\ 0 & -0.5 \\ 0 & 0.2 \\ 1 & 1 \\ -2 & -1 \\ 0.3 & 1.5 \end{pmatrix}$	$\begin{pmatrix} 5.23 & 0 \\ 0 & -2.13 \\ 0.36 & 0.19 \\ 1 & 1 \\ -0.33 & -1.09 \\ 2.70 & 1.68 \end{pmatrix}$	$\begin{pmatrix} -4.54 & 0 \\ 0 & -0.50 \\ 0 & 0.19 \\ 1 & 1 \\ -1.33 & -0.72 \\ 0.34 & 1.24 \end{pmatrix}$

Table 1. True coefficients of matrix $\bar{K}$ and identified values.

A set of noisy data of the state variables S_2, S_3, E, X, P, O and of the gaseous flow rates q_{O_2} and q_{CO_2} is produced by simulation as described in Section 2. The goal is to determine the 6×2 matrix $\bar{K}$ from this data set. More specifically, a question that we want to address is to determine, from the data, which of the two reactions produces the enzyme E.

This approach has been carried out both on the basis of a temporal analysis using moving average functions (as in Bernard and Bastin (2005b)), and on MFs. We have chosen a Fourier basis for the analysis, keeping 20 modes. Finally, using the two presented methods, we end up with two estimates of matrix $\bar{K}$ (see Table 1).

It is worth noting that the identified matrix $\bar{K}$ is much more accurate using the MFs (average error 7.5%) than with the temporal approach where the average error on each unknown coefficient was 39.5%. The value of the (theoretically zero) coefficient $\bar{k}_{31}$ is 0.01 (0.36 for the temporal approach) which can be neglected with respect to the other coefficients of $\bar{K}$. Hence, the unknown part of the structure of matrix $\bar{K}$ has been accurately recognised.

4. CONCLUSION

In this paper we have tested the advantage of using MFs for computing $\mathcal{I}mK$. It demonstrated that the computation of the number of reactions should rather be based on a temporal analysis, while the estimation of K is strongly improved using a basis of MFs. Through the studied example we have demonstrated that the MF based method can more accurately estimate the values of the PS coefficients in spite of noises due to measurements and low sampling frequency. We have chosen a Fourier basis of the MFs, but other choices are possible and their respective efficiency should be assessed.

The main result provided by the previous analysis is the determination of the variables which are substrates or products in the reactions or, in other words, the obtained signs of the entries of K.

Another expected result can be the determination of the variables which are not involved in a reaction, corresponding to zero elements in matrix K. However it is actually very unlikely that the analysis provides estimates of the elements of K which are exactly zero. The idea consist then in replacing the very small elements by zeros, and to validate the corresponding reaction network using the techniques presented in (Bernard and Bastin (2005b)). These validation techniques can also straightforwardly benefit from the filtering capacity of the MFs.

REFERENCES

G. Bastin and D. Dochain. *On-line estimation and adaptive control of bioreactors*. Elsevier, Amsterdam, 1990.

O. Bernard and G. Bastin. Identification of reaction schemes for bioprocesses: determination of an incompletely known yield matrix. *Bioprocess and Biosystem Engineering*, 27:293–302, 2005a.

O. Bernard and G.Bastin. On the estimation of the pseudo-stoichiometric matrix for mass balance modeling of biotechnological processes. *Math. Biosciences*, 193:51–77, 2005b.

O. Bernard, G. Bastin, C. Stentelaire, L. Lesage-Meessen, and M. Asther. Mass balance modelling of vanillin production from vanillic acid by cultures of the fungus *pycnoporus cinnabarinus* in bioreactors. *Biotech. Bioeng*, pages 558–571, 1999.

P. Bogaerts and A. Vande Wouwer. Sytematic generation of identifiable macroscopic reaction schemes. In *Proceedings of the 8th IFAC Conference on Computer Applications in Biotechnology (CAB8)*. Montral, Canada, 2001.

L. Chen and G. Bastin. Structural identifiability of the yield coefficients in bioprocess models when the reaction rates are unknown. *Math. Biosciences*, 132:35–67, 1996.

T. Chevalier, I. Schreiber, and J. Ross. Toward a systematic determination of complex reaction mechanisms. *J. Phys. Chem*, 97:6776 – 6787, 1993.

C. Delbès, R. Moletta, and J.-J. Godon. Bacterial and archaeal 16s rdna and 16s rrna dynamics during an acetate crisis in an anaerobic digestor ecosystem. *FEMS Microbiology Ecology*, 35:19–26, 2001.

M. Eiswirth, A. Freund, and J. Ross. *Mechanistic classification of chemical oscillators and the role of species*, volume 80 of *Advances in Chemical Physics*, chapter 1, pages 127–199. Wiley, New-York, 1991.

R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge MA, 1993.

H.A. Preisig and D.W.T. Rippin. Theory and application of the modulating function method. i: review and theory of the method and theory of the spline-type modulating functions. *Comput. chem. eng.*, 17: 1 – 16, 1993.

M. Shinbrot. On the analysis of linear and nonlinear systems. *Trans. ASME*, 79:547–552, 1957.