

# Quantitative analysis of metabolic networks and design of minimal bioreaction models

Georges Bastin

► **To cite this version:**

Georges Bastin. Quantitative analysis of metabolic networks and design of minimal bioreaction models. Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées, INRIA, 2008, 9. <hal-01277922>

**HAL Id: hal-01277922**

**<https://hal.inria.fr/hal-01277922>**

Submitted on 23 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Quantitative analysis of metabolic networks and design of minimal bioreaction models

### A brief tutorial

Georges Bastin

Département d'Ingénierie Mathématique  
Université catholique de Louvain  
1348 Louvain-la-Neuve  
Belgium  
Georges.Bastin@uclouvain.be  
www.inma.ucl.ac.be/~bastin

**ABSTRACT.** This tutorial paper is concerned with the design of macroscopic bioreaction models on the basis of a quantitative analysis of the underlying cell metabolism. The paper starts with a review of two fundamental algebraic techniques for the quantitative analysis of metabolic networks : (i) the decomposition of complex metabolic networks into elementary pathways (or elementary modes), (ii) the metabolic flux analysis which aims at computing the entire intracellular flux distribution from a limited number of flux measurements. Then it is discussed how these two fundamental techniques can be exploited to design minimal bioreaction models by using a systematic model reduction approach that automatically produces a family of equivalent minimal models which are fully compatible with the underlying metabolism and consistent with the available experimental data. The theory is illustrated with an experimental case-study on CHO cells.

**RÉSUMÉ.** Cet article tutorial traite de la conception de modèles de bioréactions macroscopiques sur la base d'une analyse quantitative du métabolisme cellulaire sous-jacent. L'article commence par un rappel de deux techniques algébriques fondamentales pour l'analyse quantitative des réseaux métaboliques : (i) la décomposition des réseaux métaboliques complexes en chemins élémentaires (ou modes élémentaires), (ii) l'analyse des flux métaboliques qui vise à calculer la totalité des flux métaboliques intracellulaires à partir d'un ensemble limité de mesures. On montre ensuite comment ces deux techniques peuvent être exploitées pour concevoir des modèles minimaux de bioréactions en utilisant une approche systématique de réduction de modèle qui produit automatiquement une famille de modèles minimaux équivalents compatibles non seulement avec les données expérimentales mais aussi avec le métabolisme sous-jacent. La théorie est illustrée avec une étude de cas expérimentale sur des cellules CHO.

**KEYWORDS :** Bioreactions, Metabolic flux analysis, Metabolic networks, Model reduction

**MOTS-CLÉS :** Analyse des flux métaboliques, Bioréactions, Réseaux métaboliques, Réduction de modèle

---

## 1. Introduction

The issue of quantitative bioprocess modelling from extracellular measurements is a central issue in bioengineering (e.g. [9]). In classical macroscopic models, the biomass is viewed as a catalyst for the conversion of substrates into products. The process is represented by a set of so-called *bioreactions* that directly connect the substrates to the products, without making an explicit reference to the intracellular metabolism.

A more recent trend is to base the design of macroscopic bioreaction models on a quantitative analysis of the underlying cell metabolism (e.g. [5], [12], [14], [20]). Our concern in this paper is to give a brief tutorial presentation of the theory that underlies this methodology. We start with a review of two fundamental techniques for the quantitative analysis of metabolic networks : (i) the decomposition of complex metabolic networks into elementary pathways (or elementary modes), (ii) the metabolic flux analysis which aims at computing the entire intracellular flux distribution from a limited number of flux measurements. Then we discuss how these two fundamental techniques can be exploited to design minimal bioreaction models by using a systematic “model reduction” approach that automatically produces a family of equivalent minimal models which are fully compatible with the underlying metabolism and consistent with the available experimental data.

As a matter of illustration and motivation to the theory, we consider the example of chinese hamster ovary (CHO) cells cultivated in batch mode in stirred flasks ([1]).

---

## 2. Metabolic networks

The intracellular metabolism of living cells is usually represented by a metabolic network under the form of a directed hypergraph that encodes a set of elementary biochemical reactions taking place within the cell. In this hypergraph, the nodes represent the involved metabolites and the edges represent the metabolic fluxes. A typical example of metabolic network is shown in Fig.2. The metabolic network involves two groups of nodes: boundary nodes and internal nodes. Boundary nodes have only either incoming or outgoing edges, but not both together. Boundary nodes are further separated into input (or initial) and output (or terminal) nodes. Input nodes correspond to substrates that are only consumed but not produced. Output nodes correspond to final products that are only produced but not consumed. In contrast, the internal (or intermediary) nodes are the nodes that have necessarily both incoming and outgoing incident edges. They correspond to intracellular metabolites that are produced by some of the metabolic reactions and consumed by other reactions inside the cell.

We adopt the quasi steady-state paradigm of metabolic flux analysis (MFA) (e.g. [17]). This means that for each internal metabolite of the network, it is assumed that the net sum of production and consumption fluxes, weighted by their stoichiometric coefficients, is zero. This is expressed by the algebraic relation:

$$\mathbf{N}\mathbf{v} = \mathbf{0} \quad \mathbf{v} \geq \mathbf{0} \quad (1)$$

where  $\mathbf{v} = (v_1, v_2, \dots, v_m)^T$  is the  $m$ -dimensional vector of fluxes and  $\mathbf{N} = [n_{ij}]$  is the  $n \times m$  stoichiometric matrix of the metabolic network ( $m$  is the number of fluxes and  $n$

the number of internal nodes of the network). More precisely, a flux  $v_j$  denotes the rate of reaction  $j$  and a non-zero  $n_{ij}$  is the stoichiometric coefficient of metabolite  $i$  in reaction  $j$ .

We also introduce the following notations :  $\mathbf{v}_s$  for the vector of the specific uptake rates of the initial substrates and  $\mathbf{v}_p$  for the vector of the specific production rates of the final products. By definition  $\mathbf{v}_s$  and  $\mathbf{v}_p$  are linear combinations of some of the metabolic fluxes. This is expressed by defining appropriate matrices  $\mathbf{N}_s$  and  $\mathbf{N}_p$  such that

$$\mathbf{v}_s = \mathbf{N}_s \mathbf{v} \quad \mathbf{v}_p = \mathbf{N}_p \mathbf{v}. \quad (2)$$

Obviously, the specific rates  $\mathbf{v}_s$  and  $\mathbf{v}_p$  are not independent since they are quantitatively related through the intracellular metabolism represented by the metabolic network. A significant outcome of the quantitative analysis is precisely to elucidate this relation more deeply.

---

### 3. Elementary modes and input/output bioreactions

For a given metabolic network, the set  $\mathcal{S}$  of possible flux distributions is the set of vectors  $\mathbf{v}$  that satisfy the set (1) of homogeneous linear equalities and inequalities. Each possible  $\mathbf{v}$  must necessarily be non-negative and belong to the kernel of the matrix  $\mathbf{N}$ . Hence the set  $\mathcal{S}$  is the pointed polyhedral cone which is the intersection of the kernel of  $\mathbf{N}$  and the nonnegative orthant. This implies that any flux distribution  $\mathbf{v}$  can be expressed as a non-negative linear combination of a set of so-called *elementary flux vectors*  $\mathbf{e}_i$  ([18]) which are the edges (or extreme rays) of the polyhedral cone and form therefore a *unique* convex basis (see e.g. [19]) of the flux space  $\mathcal{S}$ :

$$\mathbf{v} = w_1 \mathbf{e}_1 + w_2 \mathbf{e}_2 + \dots + w_p \mathbf{e}_p \quad w_i \geq 0. \quad (3)$$

The  $m \times p$  non-negative matrix  $\mathbf{E}$  with column vectors  $\mathbf{e}_i$  obviously satisfies  $\mathbf{N}\mathbf{E} = 0$  and (3) is written in matrix form as

$$\mathbf{v} = \mathbf{E}\mathbf{w} \quad \text{with} \quad \mathbf{w} \triangleq (w_1, w_2, \dots, w_p)^T. \quad (4)$$

>From a metabolic viewpoint, the vectors  $\mathbf{e}_i$  encode the simplest metabolic paths that connect the initial substrates (input nodes) to the final products (output nodes). More precisely, the non-zero entries of a basis vector  $\mathbf{e}_i$  enumerate the fluxes of a sequence of biochemical reactions starting at one or several initial substrates and ending at one or several final products. These simple pathways between initial substrates and final products are called extreme pathways (ExPa) or elementary (flux) modes (EM) of the network (e.g.[2], [9], [16]). Since the intermediate reactions are assumed to be at quasi steady-state, a single macroscopic bioreaction is then readily defined from an elementary mode by considering only the involved initial substrates and final products.

Let us now examine the relation between the specific consumption and production rates  $\mathbf{v}_s$  and  $\mathbf{v}_p$  induced by the metabolic network. From (2) and (4) it follows that

$$\begin{pmatrix} -\mathbf{v}_s \\ \mathbf{v}_p \end{pmatrix} = \begin{pmatrix} -\mathbf{N}_s \\ \mathbf{N}_p \end{pmatrix} \mathbf{E}\mathbf{w} = \mathbf{K}_e \mathbf{w} \quad (5)$$

where

$$\mathbf{K}_e \triangleq \begin{pmatrix} -\mathbf{N}_s \\ \mathbf{N}_p \end{pmatrix} \mathbf{E}$$

is the stoichiometric matrix of the set of input/output bioreactions encoded by the EMs. Hence  $\mathbf{v}_s$  and  $\mathbf{v}_p$  can be regarded as the specific uptake and production rates of a bioprocess governed by the bioreactions with stoichiometry  $\mathbf{K}_e$  and specific reaction rates  $\mathbf{w}$ . In other terms, each weighting coefficient  $w_i$  in (3) can equally be interpreted as the specific reaction rate of the bioreaction encoded by the EM  $\mathbf{e}_i$ : the flux vector  $\mathbf{v}$  is thus a linear combination of EMs whose non-negative weights are the macroscopic bioreaction rates  $w_i$ .

An important issue concerns the number of distinct EMs that are generated when computing the EMs. It may become very large because it combinatorially increases with the size of the underlying metabolic network<sup>1</sup>. Furthermore, even when the number of EMs is rather limited, it appears that the resulting set of bioreactions can be significantly redundant for the design of models that fully explain the available experimental data. There is therefore clearly an interest for reducing the model size as much as possible and trying to determine a minimal subset of bioreactions that is sufficient to build the bioreaction model (5).

---

## 4. Metabolic flux analysis

Metabolic flux analysis (MFA) is the exercise of calculating the admissible flux distributions  $\mathbf{v}$  that satisfy the steady state balance equation  $\mathbf{N}\mathbf{v} = \mathbf{0}$  together with an additional set of linear constraints added by using experimental measurements. Here we consider the case where the measurements are collected in a vector  $\mathbf{v}_m$  which is a linear function of the unknown flux distribution  $\mathbf{v}$  and is expressed as

$$\mathbf{v}_m = \mathbf{P}\mathbf{v} \quad (6)$$

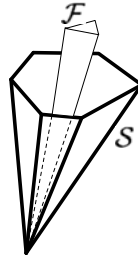
with  $\mathbf{P}$  being a known  $\dim\{\mathbf{v}_m\} \times n$  matrix. In addition, it is assumed that  $\mathbf{P}\mathbf{e}_i \neq \mathbf{0} \forall i$  or, in other terms, that the elementary flux vectors  $\mathbf{e}_i$  do not belong to the kernel of the matrix  $\mathbf{P}$ . Then, from equations (1)-(6), we have the following fundamental equation of metabolic flux analysis

$$\Sigma \begin{pmatrix} \mathbf{v} \\ 1 \end{pmatrix} = \mathbf{0} \quad \text{with} \quad \Sigma \triangleq \begin{pmatrix} \mathbf{N} & \mathbf{0} \\ \mathbf{P} & -\mathbf{v}_m \end{pmatrix} \quad \text{and} \quad \mathbf{v} \geq \mathbf{0}. \quad (7)$$

For a given metabolic network and a given set of measurements, the solution of the MFA problem is defined as the space  $\mathcal{F}$  of admissible flux distributions. i.e. the set of non-negative vectors  $\mathbf{v}$  that satisfy the finite set (7) of homogeneous linear equalities and inequalities. Each admissible  $\mathbf{v}$  must be such that the non-negative vector  $(\mathbf{v}^T \ 1)$  belongs

---

1. The Double Description (DD) method ([8]) is the simplest known algorithm for computing the convex basis of the solution space (see [3] for a review). In the context of metabolic networks various refinements have been proposed that differ from the original DD algorithm mainly by their initialization. A first specific algorithm was presented by [15]. Recently, the implementation of the DD method for metabolic networks has received various further improvements (e.g. [4] and [6]).



**Figure 1.** Illustration of the flux spaces  $S$  and  $\mathcal{F}$ .

to the kernel of the matrix  $\Sigma$ . Hence, as emphasized in [11, Chapter 4]-[13], the set  $\mathcal{F}$  is also a pointed polyhedral cone in the positive orthant  $\mathbb{R}_+^m$ . This means that any admissible flux distribution  $\mathbf{v}$  can be expressed as a convex combination of a set of  $q$  non-negative basis vectors  $\mathbf{f}_i$  which are the edges (or extreme rays) of this polyhedral cone and form therefore a *unique* convex basis of the flux space  $\mathcal{F}$ . In other words, the solution of the MFA problem is the *admissible flux space*  $\mathcal{F}$  defined as

$$\mathcal{F} \triangleq \left\{ \mathbf{v} : \mathbf{v} = \sum_{i=1}^q \omega_i \mathbf{f}_i, \omega_i \geq 0, \sum_{i=1}^q \omega_i = 1 \right\}. \quad (8)$$

The admissible flux space  $\mathcal{F}$  is a subset of the possible flux space  $S$  generated by the elementary modes. In geometric terms, the pointed cone  $\mathcal{F}$  is a subcone of the pointed cone  $S$  as illustrated in Fig.1. The smallest “hyper-rectangular” set that encloses  $\mathcal{F}$  in  $\mathbb{R}^m$  is called the *flux spectrum* (e.g. [7]) and is defined as the set

$$\mathcal{F}_o = \left\{ \mathbf{v} : v_i^{\min} \leq v_i \leq v_i^{\max} \right\}$$

where the bounds  $v_i^{\min}$  and  $v_i^{\max}$  are defined from the convex basis vectors as follows:

$$v_i^{\min} \triangleq \min \{ f_{ki}, k = 1, \dots, p \}, \quad v_i^{\max} \triangleq \max \{ f_{ki}, k = 1, \dots, p \},$$

where  $f_{ki}$  denotes the  $i$ -th element of the basis vector  $\mathbf{f}_k$ .

## 5. Minimal bioreaction models based on minimal sets of elementary modes

For any admissible flux vector  $\mathbf{v}$  in the cone  $\mathcal{F}$  satisfying equation (7), it must be emphasized that *the decomposition of  $\mathbf{v}$  in the convex basis  $\{\mathbf{e}_i\}$  is not unique*. As we shall see, this is the algebraic expression of the fact that the set of bioreactions that may be used to compute  $\mathbf{v}$  with equation (4) or  $\mathbf{v}_s$  and  $\mathbf{v}_p$  with equation (5) is redundant. Using (4), system (7) is equivalent to the system:

$$\begin{pmatrix} \mathbf{NE} \\ \mathbf{PE} \end{pmatrix} \mathbf{w} = \begin{pmatrix} \mathbf{0} \\ \mathbf{v}_m \end{pmatrix} \quad \mathbf{w} \geq 0. \quad (9)$$

We observe that the first equation  $\mathbf{NEw} = \mathbf{0}$  is trivially satisfied independently of  $\mathbf{w}$  since  $\mathbf{NE} = \mathbf{0}$  by definition. Hence, system (9) may be reduced to the second equation:

$$\mathbf{PEw} = \mathbf{v}_m \quad \mathbf{w} \geq 0.$$

or equivalently:

$$\begin{pmatrix} \mathbf{PE} & -\mathbf{v}_m \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ 1 \end{pmatrix} = 0 \quad \mathbf{w} \geq 0. \quad (10)$$

In this form, it is clear that the set of admissible reaction rate vectors  $\mathbf{w}$  that satisfy (10) again constitutes a convex polyhedral cone. Therefore there exists a set of appropriate edge vectors  $\mathbf{h}_i$  such that any arbitrary convex combination of the form:

$$\mathbf{w} = \sum_i z_i \mathbf{h}_i \quad z_i \geq 0 \quad \sum_i z_i = 1 \quad (11)$$

is necessarily an admissible  $\mathbf{w}$  satisfying (10). The convex basis vectors  $\mathbf{h}_i$  have an important and critical property : the number of non-zero entries is **at most** equal to the size of the vector  $\mathbf{v}_m$  i.e. the number of measurements (see [3] and Section 3.5 in [11]). From a metabolic viewpoint, each vector  $\mathbf{h}_i$  is a particular solution  $\mathbf{w}$  of (10), or equivalently a particular way (among an infinity) of computing an admissible flux distribution  $\mathbf{v}_i$ :

$$\mathbf{v}_i = \mathbf{E}\mathbf{h}_i \quad \mathbf{v}_i \in \mathcal{F} \quad (12)$$

Of course in this expression, the non-zero entries of the vector  $\mathbf{h}_i$  are interpreted as the weights of the respective contributions of the corresponding EMs in the computation of the flux distribution  $\mathbf{v}_i$ . But, at the same time, they can also be interpreted as being the specific rates of the bioreactions that are encoded by the EMs and are involved in the bioreaction model (5).

Hence each convex basis vector  $\mathbf{h}_i$  brings two different pieces of information. First it tells which EMs and consequently which input/output bioreactions are sufficient to compute internal flux distributions that are consistent with the measurements  $\mathbf{v}_m$ . These EMs are designated by the position of the non-zero entries of  $\mathbf{h}_i$ . Secondly, the value of each non-zero entry of  $\mathbf{h}_i$  is the value of the reaction rate of the corresponding bioreaction.

For each basis vector  $\mathbf{h}_i$ , we can then define a selection matrix  $\mathbf{S}_i$  that encodes the corresponding selection of bioreactions. Then equation (5) is reduced to a minimal form:

$$\begin{pmatrix} -\mathbf{v}_s \\ \mathbf{v}_p \end{pmatrix} = \mathbf{K}_i \mathbf{r}_i \quad (13)$$

where  $\mathbf{K}_i \triangleq \mathbf{K}_e \mathbf{S}_i$  and  $\mathbf{r}_i \triangleq (\mathbf{S}_i)^T \mathbf{h}_i$  respectively denote the stoichiometric matrix and the vector of the specific reaction rates of the selected minimal set of bioreactions.

Therefore, we see that the computation of the convex basis vectors  $\mathbf{h}_i$  provides the tool for determining all the minimal input/output bioreaction models that are both consistent with the intracellular metabolism and the experimental measurements.

---

## 6. Case-study

As a matter of illustration and motivation to the methodology presented above, we consider the example of chinese hamster ovary (CHO) cells cultivated in batch mode in stirred flasks in a serum-free medium ([1]). During the growth phase, the cell metabolism is described by the metabolic network presented in Fig.2. This network describes only the part of the metabolism concerned with the utilisation of the two main energetic nutrients (glucose and glutamine). The metabolism of the amino acids provided by the

culture medium is not considered. The network involves the Glycolysis pathway, the Pentose-Phosphate pathway and the Krebs cycle. Moreover it is assumed that a part of the glutamine is used for the making of nucleotides which are lumped into a single species with equal shares of purines and pyrimidines (see [11] and [14] for further motivation and details).

In this network, there are

- two input substrates : Glucose and Glutamine;
- five output products : Lactate, CO<sub>2</sub>, NH<sub>4</sub>, Alanine and Nucleotides;
- $n = 18$  internal metabolites : Glucose-6-phosphate, Fructose-6-Phosphate, Dihydroxy-acetone-phosphate, Glyceraldehyde-3 phosphate, Pyruvate, Acetyl-coA, Citrate,  $\alpha$ -ketoglutarate, Fumarate, Malate, Oxaloacetate, Aspartate, Glutamate, CO<sub>2</sub>, Ribose-5-Phosphate, Ribulose-5-Phosphate, Xylose-5-Phosphate, Erythrose-4-Phosphate;
- $m = 24$  metabolic fluxes denoted  $v_1$  to  $v_{24}$  in Fig. 2.

Without loss of generality, all the intermediate metabolites that are not located at branch points have been omitted from the network. The stoichiometric matrix  $\mathbf{N}$  is given in Table 1. The matrices  $\mathbf{N}_s$  and  $\mathbf{N}_p$  are given in Tables 2 and 3 respectively.

The network has eleven elementary modes given in Table 6 (computed with META-TOOL<sup>2</sup>) and from which the following set of input/output bioreactions can be derived:

- (e<sub>1</sub>) Glucose  $\rightarrow$  2 Lactate
- (e<sub>2</sub>) 2 Glucose + 3 Glutamine  $\rightarrow$  Alanine + Nucleotide + 9 CO<sub>2</sub>
- (e<sub>3</sub>) Glutamine  $\rightarrow$  Lactate + 2 NH<sub>4</sub> + 2 CO<sub>2</sub>
- (e<sub>4</sub>) Glutamine  $\rightarrow$  2 NH<sub>4</sub> + 5 CO<sub>2</sub>
- (e<sub>5</sub>) Glutamine  $\rightarrow$  Alanine + NH<sub>4</sub> + 2 CO<sub>2</sub>
- (e<sub>6</sub>) 2 Glucose + 3 Glutamine  $\rightarrow$  Lactate + Alanine + Nucleotide + 6 CO<sub>2</sub>
- (e<sub>7</sub>) 3 Glucose  $\rightarrow$  5 Lactate + 3 CO<sub>2</sub>
- (e<sub>8</sub>) 2 Glucose + 3 Glutamine  $\rightarrow$  2 Lactate + Nucleotide + NH<sub>4</sub> + 6CO<sub>2</sub>
- (e<sub>9</sub>) Glucose  $\rightarrow$  6 CO<sub>2</sub>
- (e<sub>10</sub>) Glucose  $\rightarrow$  6 CO<sub>2</sub>
- (e<sub>11</sub>) 2 Glucose + 3 Glutamine  $\rightarrow$  Nucleotide + NH<sub>4</sub> + 12 CO<sub>2</sub>

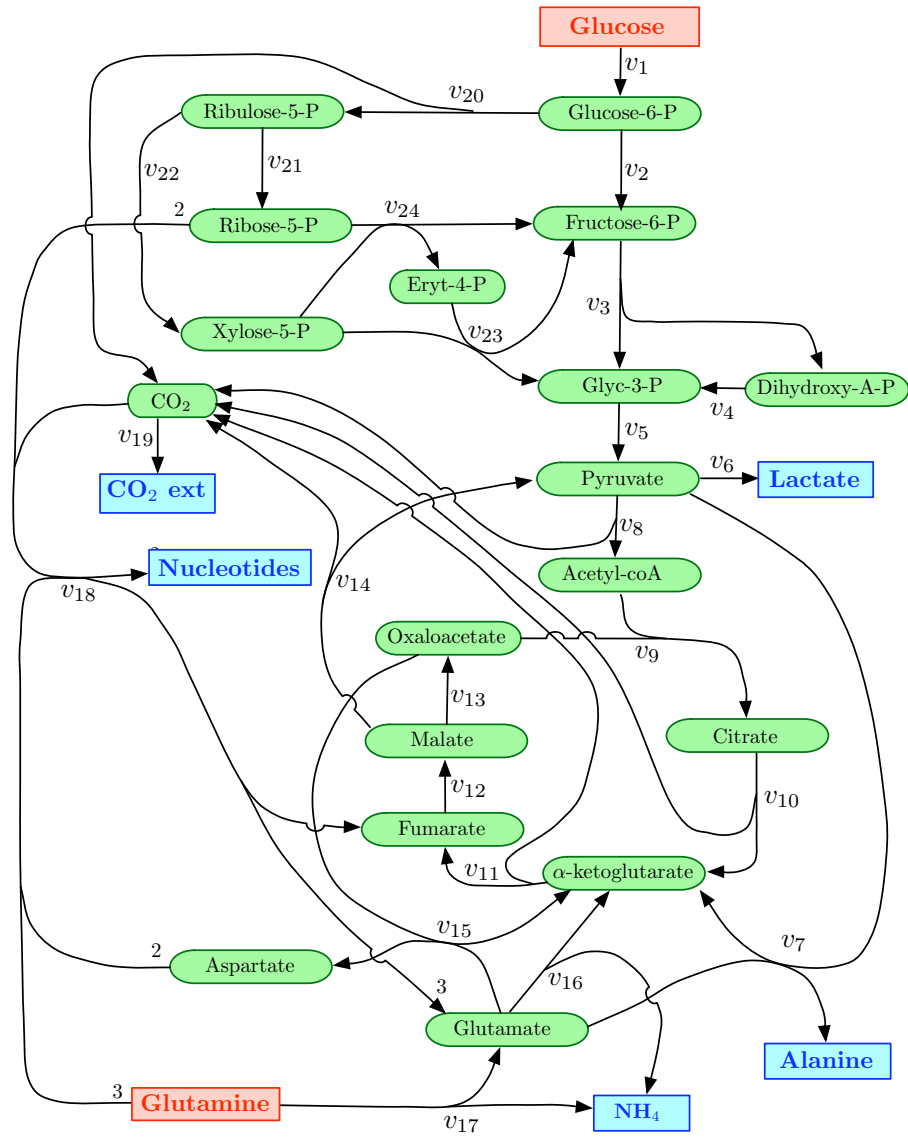
We observe that the two bioreactions corresponding to elementary modes e<sub>9</sub> and e<sub>10</sub> are identical (Glucose  $\rightarrow$  6 CO<sub>2</sub>) although the two concerned elementary pathways are different.

It can be checked that the stoichiometric matrix of this set of bioreactions is given by the matrix product

$$\mathbf{K}_e \triangleq \begin{pmatrix} -\mathbf{N}_s \\ \mathbf{N}_p \end{pmatrix} \mathbf{E}.$$

2. <http://pinguin.biologie.uni-jena.de/bioinformatik/networks/> (see also [10]).





**Figure 2.** Metabolic network : rectangular boxes represent input/output nodes, elliptic boxes represent internal nodes. (The numbers along some arrows indicate stoichiometric coefficients).

Moreover, there are five measured extra-cellular species : the two substrates (Glucose and Glutamine) and three excreted products (Lactate, Ammonia, Alanine). The values of the average specific uptake and excretion rates (vector  $\mathbf{v}_m$ ), computed by linear regression during the growth phase (see [12]), are given in Table 4. The corresponding matrix  $\mathbf{P}$  is given in Table 5. We observe that in this case:

the matrix  $\mathbf{P}$  is a sub-matrix of  $\begin{pmatrix} \mathbf{N}_s \\ \mathbf{N}_p \end{pmatrix}$ .

The admissible flux space  $\mathcal{F}$  is generated by a convex basis that includes  $p = 2$  basis vectors that are given in Table 7 (computed with METATOOL). Obviously the values given in this Table are also the limiting values  $v_i^{\min}$  and  $v_i^{\max}$  of the flux spectrum. It is remarkable that, although the MFA problem is here underdetermined, the values of the fluxes  $v_1, v_6, v_7, v_{14}, v_{15}, v_{16}, v_{17}, v_{18}$  and  $v_{19}$  are exactly given without uncertainty. This is obviously normal for the three fluxes  $v_1$  (Glucose),  $v_6$  (Lactate),  $v_7$  (Alanine) that are constrained to be equal to their measured values. But we observe that it is also the case for other fluxes like for instance the production fluxes of Nucleotides  $v_{18}$  and  $\text{CO}_2$   $v_{19}$  that are not measured at all, and also for some intracellular fluxes like for instance the anaplerotic flux  $v_{14}$ .

We then compute the set of vectors  $\mathbf{h}_i$  and the result is shown in Table 8. We observe that there are 24 different vectors  $\mathbf{h}_i$  in this Table. They all produce an admissible flux distribution  $\mathbf{v}_i \in \mathcal{F}$  when premultiplied by the matrix  $\mathbf{E}$  as expected according to (12). Furthermore, as predicted by the theory, we also observe that there are exactly 5 non-zero entries in each vector  $\mathbf{h}_i$ . From these observations, we can conclude that there are 24 different equivalent minimal bioreaction models of the form (13) for the considered process. For each of these models, Table 8 tells us which 5 bioreactions (among the eleven) are used and the value of their reaction rates. As it can be concluded from Table 7, all these minimal models are **equivalent** because they all provide exactly the same values of  $\mathbf{v}_s$  and  $\mathbf{v}_p$ , not only for the measured species but also for the species that are not measured.

---

## 7. References

- [1] J.S. Ballez, J. Mols, J. Burteau, S.N. Agathos, and Y.-J. Schneider. Plant protein hydrolysates support CHO-320 cells proliferation and recombinant ifn-gamma production in suspension and inside microcarriers in protein-free media. *Cytotechnology*, 44:103-114, 2004.
- [2] S.L. Bell and B.O. Palsson. EXPA: a program for calculating extreme pathways in biochemical reaction networks. *Bioinformatics*, 21(8):1739-1740, 2005.
- [3] K. Fukuda and A. Prodon. The double description method revisited. In R. Euler and M. E. Deza I. Manoussakis, editors, *Combinatorics and Computer Science*, volume 1120 of *Lecture Notes in Computation Sciences*, pages 91-111. Springer-Verlag, 1996.
- [4] J. Gagneur and S. Klamt. Computation of elementary modes : a unifying framework and the new binary approach. *BMC Bioinformatics*, 5:175, 2004.
- [5] J.E. Haag, A. Vande Wouwer, and P. Bogaerts. Dynamic modeling of complex biological systems : a link between metabolic and macroscopic descriptions. *Mathematical Biosciences*, 193:25-49, 2005.
- [6] S. Klamt, J. Gagneur, and A. von Kamp. Algorithmic approaches for computing elementary modes in large biochemical networks. *IEE Proceedings Systems Biology*, 152:249-255, 2005.

- [7] F. Llaneras and J. Pico. An interval approach for dealing with flux distributions and elementary modes activity patterns. *Journal of Theoretical Biology*, 246:290-308, 2007.
- [8] T.S. Motzkin, H. Raiffa, G.L. Thompson, and R.M. Thrall. The double description method. In H.W. Kuhn and A.W. Tucker, editors, *Contribution to the Theory of Games Vol. II*, volume 28 of *Annals of Mathematical Studies*, pages 51-73, Princeton, New Jersey, 1953. Princeton University Press.
- [9] J. Nielsen, J. Villadsen, and G. Liden. *Bioreaction Engineering Principles*. Kluwer Academic, 2002.
- [10] T. Pfeiffer, I. Sanchez-Valdenebro, J.C. Nu, F. Montero, and S. Schuster. Metatool : for studying metabolic networks. *Bioinformatics*, 15(3):251-257, March 1999.
- [11] A. Provost. *Metabolic design of dynamic bioreaction models*. PhD thesis, Faculty of Engineering, Université catholique de Louvain, November 2006.
- [12] A. Provost and G. Bastin. Dynamical metabolic modelling under the balanced growth condition. *Journal of Process Control*, 14(7):717-728, 2004.
- [13] A. Provost and G. Bastin. Metabolic flux analysis : an approach for solving non-stationary underdetermined systems. In *CD-Rom Proceedings 5th MATHMOD Conference*, Paper 207 in Session SP33, Vienna, Austria, 2006.
- [14] A. Provost, G. Bastin, S.N. Agathos, and Y-J. Schneider. Metabolic design of macroscopic bioreaction models : Application to Chinese hamster ovary cells. *Bioprocess and Biosystems Engineering*, 29(5-6):349-366, 2006.
- [15] R. Schuster and S. Schuster. RePned algorithm and computer program for calculating all non-negative fluxes admissible in steady states of biochemical reaction systems with or without some flux rates fixed. *Computer Applications in the Biosciences*, 9(1):79-85, February 1993.
- [16] S. Schuster, D.A. Fell, and T. Dandekar. Detection of elementary flux modes in biochemical networks : a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, 17(2):53-60, February 1999.
- [17] G. Stephanopoulos, J. Nielsen, and A. Aristidou. *Metabolic Engineering : Principles and Methodologies*. Academic Press, San Diego, 1998.
- [18] R. Urbanczik. Enumerating constrained elementary flux vectors of metabolic networks. *IET Systems Biology*, 1(5):274-279, 2007.
- [19] H. Weyl. The elementary theory of convex polyhedra. In *Contributions to the Theory of Games Vol. I*, *Annals of Mathematical Studies*, pages 3-18, Princeton, New Jersey, 1950. Princeton University Press.
- [20] F. Zhou, J-X. Bi, A-P. Zeng, and J-Q. Yuan. A macrokinetic and regulator model for myeloma cell culture based on metabolic balance of pathways. *Process Biochemistry*, 41:2207-2217, 2006.

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$	$v_9$	$v_{10}$	$v_{11}$	$v_{12}$
Glucose-6-P	1	-1	0	0	0	0	0	0	0	0	0	0
Fructose-6-P	0	1	-1	0	0	0	0	0	0	0	0	0
Glyc-3-P	0	0	1	1	-1	0	0	0	0	0	0	0
Dihydroxy-A-P	0	0	1	-1	0	0	0	0	0	0	0	0
Pyruvate	0	0	0	0	1	-1	-1	-1	0	0	0	0
Acetyl-coA	0	0	0	0	0	0	0	1	-1	0	0	0
Citrate	0	0	0	0	0	0	0	0	1	-1	0	0
$\alpha$ -ketoglutarate	0	0	0	0	0	0	1	0	0	1	-1	0
Fumarate	0	0	0	0	0	0	0	0	0	0	1	-1
Malate	0	0	0	0	0	0	0	0	0	0	0	1
Oxaloacetate	0	0	0	0	0	0	0	0	-1	0	0	0
Glutamate	0	0	0	0	0	0	-1	0	0	0	0	0
Aspartate	0	0	0	0	0	0	0	0	0	0	0	0
Ribulose-5-P	0	0	0	0	0	0	0	0	0	0	0	0
Ribose-5-P	0	0	0	0	0	0	0	0	0	0	0	0
Xylose-5-P	0	0	0	0	0	0	0	0	0	0	0	0
Erythrose-4-P	0	0	0	0	0	0	0	0	0	0	0	0
CO <sub>2</sub>	0	0	0	0	0	0	0	1	0	1	1	0

	$v_{13}$	$v_{14}$	$v_{15}$	$v_{16}$	$v_{17}$	$v_{18}$	$v_{19}$	$v_{20}$	$v_{21}$	$v_{22}$	$v_{23}$	$v_{24}$
Glucose-6-P	0	0	0	0	0	0	0	-1	0	0	0	0
Fructose-6-P	0	0	0	0	0	0	0	0	0	0	1	1
Glyc-3-P	0	0	0	0	0	0	0	0	0	0	1	0
Dihydroxy-A-P	0	0	0	0	0	0	0	0	0	0	0	0
Pyruvate	0	1	0	0	0	0	0	0	0	0	0	0
Acetyl-coA	0	0	0	0	0	0	0	0	0	0	0	0
Citrate	0	0	0	0	0	0	0	0	0	0	0	0
$\alpha$ -ketoglutarate	0	0	1	1	0	0	0	0	0	0	0	0
Fumarate	0	0	0	0	0	1	0	0	0	0	0	0
Malate	-1	-1	0	0	0	0	0	0	0	0	0	0
Oxaloacetate	1	0	-1	0	0	0	0	0	0	0	0	0
Glutamate	0	0	-1	-1	1	3	0	0	0	0	0	0
Aspartate	0	0	1	0	0	-2	0	0	0	0	0	0
Ribulose-5-P	0	0	0	0	0	0	0	1	-1	-1	0	0
Ribose-5-P	0	0	0	0	0	-2	0	0	1	0	0	-1
Xylose-5-P	0	0	0	0	0	0	0	0	0	1	-1	-1
Erythrose-4-P	0	0	0	0	0	0	0	0	0	0	-1	1
CO <sub>2</sub>	0	1	0	0	0	0	-1	1	0	0	0	0

Table 1. Stoichiometric Matrix N

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$	$v_9$	$v_{10}$	$v_{11}$	$v_{12}$
Glucose	1	0	0	0	0	0	0	0	0	0	0	0
Glutamine	0	0	0	0	0	0	0	0	0	0	0	0

	$v_{13}$	$v_{14}$	$v_{15}$	$v_{16}$	$v_{17}$	$v_{18}$	$v_{19}$	$v_{20}$	$v_{21}$	$v_{22}$	$v_{23}$	$v_{24}$
Glucose	0	0	0	0	0	0	0	0	0	0	0	0
Glutamine	0	0	0	0	1	3	0	0	0	0	0	0

**Table 2.** Matrix  $N_s$

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$	$v_9$	$v_{10}$	$v_{11}$	$v_{12}$
Lactate	0	0	0	0	0	1	0	0	0	0	0	0
NH <sub>4</sub>	0	0	0	0	0	0	0	0	0	0	0	0
Alanine	0	0	0	0	0	0	1	0	0	0	0	0
CO <sub>2</sub>	0	0	0	0	0	0	0	0	0	0	0	0
Nucleotides	0	0	0	0	0	0	0	0	0	0	0	0

	$v_{13}$	$v_{14}$	$v_{15}$	$v_{16}$	$v_{17}$	$v_{18}$	$v_{19}$	$v_{20}$	$v_{21}$	$v_{22}$	$v_{23}$	$v_{24}$
Lactate	0	0	0	0	0	0	0	0	0	0	0	0
NH <sub>4</sub>	0	0	0	1	1	0	0	0	0	0	0	0
Alanine	0	0	0	0	0	0	0	0	0	0	0	0
CO <sub>2</sub> ext	0	0	0	0	0	0	1	0	0	0	0	0
Nucleotides	0	0	0	0	0	1	0	0	0	0	0	0

**Table 3.** Matrix  $N_p$

	Glucose	Glutamine	Lactate	NH <sub>4</sub>	Alanine
$v_m$	4.0546	1.1860	7.3949	0.9617	0.2686

**Table 4.** Specific uptake and excretion rates (mM/(d×10<sup>9</sup>cells)).

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$	$v_9$	$v_{10}$	$v_{11}$	$v_{12}$
Glucose	1	0	0	0	0	0	0	0	0	0	0	0
Glutamine	0	0	0	0	0	0	0	0	0	0	0	0
Lactate	0	0	0	0	0	1	0	0	0	0	0	0
NH <sub>4</sub>	0	0	0	0	0	0	0	0	0	0	0	0
Alanine	0	0	0	0	0	0	1	0	0	0	0	0

	$v_{13}$	$v_{14}$	$v_{15}$	$v_{16}$	$v_{17}$	$v_{18}$	$v_{19}$	$v_{20}$	$v_{21}$	$v_{22}$	$v_{23}$	$v_{24}$
Glucose	0	0	0	0	0	0	0	0	0	0	0	0
Glutamine	0	0	0	0	1	3	0	0	0	0	0	0
Lactate	0	0	0	0	0	0	0	0	0	0	0	0
NH <sub>4</sub>	0	0	0	1	1	0	0	0	0	0	0	0
Alanine	0	0	0	0	0	0	0	0	0	0	0	0

**Table 5.** Measurement matrix  $P$

	e <sub>1</sub>	e <sub>2</sub>	e <sub>3</sub>	e <sub>4</sub>	e <sub>5</sub>	e <sub>6</sub>	e <sub>7</sub>	e <sub>8</sub>	e <sub>9</sub>	e <sub>10</sub>	e <sub>11</sub>
v <sub>1</sub>	1	2	0	0	0	2	3	2	1	3	2
v <sub>2</sub>	1	0	0	0	0	0	0	0	1	0	0
v <sub>3</sub>	1	0	0	0	0	0	2	0	1	2	0
v <sub>4</sub>	1	0	0	0	0	0	2	0	1	2	0
v <sub>5</sub>	2	0	0	0	0	0	5	0	2	5	0
v <sub>6</sub>	2	0	1	0	0	1	5	2	0	0	0
v <sub>7</sub>	0	1	0	0	1	1	0	0	0	0	0
v <sub>8</sub>	0	1	0	1	0	0	0	0	2	5	2
v <sub>9</sub>	0	1	0	1	0	0	0	0	2	5	2
v <sub>10</sub>	0	1	0	1	0	0	0	0	2	5	2
v <sub>11</sub>	0	4	1	2	1	3	0	3	2	5	5
v <sub>12</sub>	0	5	1	2	1	4	0	4	2	5	6
v <sub>13</sub>	0	3	0	1	0	2	0	2	2	5	4
v <sub>14</sub>	0	2	1	1	1	2	0	2	0	0	2
v <sub>15</sub>	0	2	0	0	0	2	0	2	0	0	2
v <sub>16</sub>	0	0	1	1	0	0	0	1	0	0	1
v <sub>17</sub>	0	0	1	1	1	0	0	0	0	0	0
v <sub>18</sub>	0	1	0	0	0	1	0	1	0	0	1
v <sub>19</sub>	0	9	2	5	2	6	3	6	6	18	12
v <sub>20</sub>	0	2	0	0	0	2	3	2	0	3	2
v <sub>21</sub>	0	2	0	0	0	2	1	2	0	1	2
v <sub>22</sub>	0	0	0	0	0	0	2	0	0	2	0
v <sub>23</sub>	0	0	0	0	0	0	1	0	0	1	0
v <sub>24</sub>	0	0	0	0	0	0	1	0	0	1	0

**Table 6.** Matrix **E** of elementary modes.

	$f_1$	$f_2$
$v_1$	4.0546	4.0546
$v_2$	3.5979	2.1279
$v_3$	3.5979	3.1079
$v_4$	3.5979	3.1079
$v_5$	7.1958	6.7058
$v_6$	7.3949	7.3949
$v_7$	0.2686	0.2686
$v_8$	0.4900	0.0
$v_9$	0.4900	0.0
$v_{10}$	0.4900	0.0
$v_{11}$	1.6760	1.1860
$v_{12}$	1.9043	1.4143
$v_{13}$	0.9467	0.4567
$v_{14}$	0.9577	0.9577
$v_{15}$	0.4567	0.4567
$v_{16}$	0.4607	0.4607
$v_{17}$	0.5010	0.5010
$v_{18}$	0.2283	0.2283
$v_{19}$	3.8420	3.8420
$v_{20}$	0.4567	1.9267
$v_{21}$	0.4567	0.9467
$v_{22}$	0.0	0.9800
$v_{23}$	0.0	0.4900
$v_{24}$	0.0	0.4900

**Table 7.** Convex basis of the flux space (mM/(d $\times$ 10<sup>9</sup> cells))

	$h_1$	$h_2$	$h_3$	$h_4$	$h_5$	$h_6$	$h_7$	$h_8$
$e_1$	3.5833	3.4671	3.5833	3.5813	3.4671	3.5101	3.5979	3.5813
$e_2$	0	0.2283	0	0	0.2280	0	0.2283	0
$e_3$	0	0.4607	0	0.2324	0.4607	0	0.1991	0.2324
$e_4$	0.4607	0	0.4607	0	0	0.4607	0.2617	0
$e_5$	0.0403	0.0403	0.0403	0.2686	0.0403	0.0403	0.0403	0.2686
$e_6$	0.2283	0	0.2283	0	0	0.2283	0	0
$e_7$	0	0	0	0	0	0.0293	0	0
$e_8$	0	0	0	0	0	0	0	0
$e_9$	0	0	0.0146	0	0.1308	0	0	0.0167
$e_{10}$	0.0049	0.0436	0	0.0056	0	0	0	0
$e_{11}$	0	0	0	0.2283	0	0	0	0.2283

	$h_9$	$h_{10}$	$h_{11}$	$h_{12}$	$h_{13}$	$h_{14}$	$h_{15}$	$h_{16}$
$e_1$	3.3529	3.3529	3.3529	2.8129	3.4980	2.1279	3.3529	2.1279
$e_2$	0	0	0	0.2283	0	0	0	0
$e_3$	0.2324	0.2324	0.4607	0.4607	0.2324	0.2324	0.4607	0.4607
$e_4$	0	0	0	0	0	0	0	0
$e_5$	0.2686	0.2686	0.0403	0.0403	0.2686	0.2686	0.0403	0.0403
$e_6$	0	0	0.2283	0	0	0	0.2283	0.2283
$e_7$	0	0	0	0.2617	0.0333	0.4900	0	0.4900
$e_8$	0.2283	0.2283	0	0	0	0.2283	0	0
$e_9$	0	0.2450	0.2450	0	0	0	0	0
$e_{10}$	0.0817	0	0	0	0	0	0.0817	0
$e_{11}$	0	0	0	0	0.2283	0	0	0

	$h_{17}$	$h_{18}$	$h_{19}$	$h_{20}$	$h_{21}$	$h_{22}$	$h_{23}$	$h_{24}$
$e_1$	3.5979	3.5979	3.5979	3.5979	3.5979	2.8251	3.4691	3.4691
$e_2$	0.1288	0	0	0	0.0293	0	0	0
$e_3$	0	0	0.1991	0	0	0	0	0
$e_4$	0.3612	0.2324	0.0333	0.4314	0.4607	0.2324	0.2324	0.2324
$e_5$	0.1398	0.2686	0.2686	0.0695	0.0403	0.2686	0.2686	0.2686
$e_6$	0	0	0	0.1991	0.1991	0	0	0
$e_7$	0	0	0	0	0	0.2576	0	0
$e_8$	0.0995	0.0995	0	0	0	0.2283	0.2283	0.2283
$e_9$	0	0	0	0	0	0	0.1288	0
$e_{10}$	0	0	0	0	0	0	0	0.0429
$e_{11}$	0	0.1288	0.2283	0.0293	0	0	0	0

**Table 8.** The set of basis vectors  $h_i$  (mM/(d $\times$ 10<sup>9</sup> cells))