

Review

How to Tackle Underdeterminacy in Metabolic Flux Analysis? A Tutorial and Critical Review

Philippe Bogaerts ^{1,*}  and Alain Vande Wouwer ^{2,*} 

¹ 3BIO-BioControl, Université Libre de Bruxelles, Avenue F.D. Roosevelt, 50-CP 165/61, 1050 Brussels, Belgium

² Systems, Estimation, Control and Optimization (SECO), University of Mons, 7000 Mons, Belgium

* Correspondence: philippe.bogaerts@ulb.be (P.B.); alain.vandewouwer@umons.ac.be (A.V.W.)

† These authors contributed equally to this work.

Abstract: Metabolic flux analysis is often (not to say almost always) faced with system underdeterminacy. Indeed, the linear algebraic system formed by the steady-state mass balance equations around the intracellular metabolites and the equality constraints related to the measurements of extracellular fluxes do not define a unique solution for the distribution of intracellular fluxes, but instead a set of solutions belonging to a convex polytope. Various methods have been proposed to tackle this underdeterminacy, including flux pathway analysis, flux balance analysis, flux variability analysis and sampling. These approaches are reviewed in this article and a toy example supports the discussion with illustrative numerical results.

Keywords: flux variability analysis; flux balance analysis; sampling; metabolic network; elementary flux modes



Citation: Bogaerts, P.;

Vande Wouwer, A. How to Tackle Underdeterminacy in Metabolic Flux Analysis? A Tutorial and Critical Review. *Processes* **2021**, *9*, 1577. <https://doi.org/10.3390/pr9091577>

Academic Editor: Andrzej Pawłowski

Received: 31 July 2021

Accepted: 27 August 2021

Published: 2 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Computational approaches for studying the flux distribution inside metabolic networks of microbial strains or mammalian cell lines have gained a tremendous importance in biotechnology. Indeed, the production of high-added value biochemicals is based on large-scale cultures of genetically engineered strains, and the determination of the flux distribution provides insight into the biosynthesis pathways, the impact of metabolic engineering and the influence of the culture conditions. Different approaches have been developed to compute this flux distribution, which are based on a common assumption that the intracellular metabolites do not accumulate, or in other words, that the cell is in a metabolic pseudo-steady state [1]. This assumption leads to a system of mass-balance equations of the form:

$$N\underline{v} = 0 \quad v_i \geq 0 \quad \forall i \quad (1)$$

where $N \in \mathbb{R}^{n_s \times n_v}$ is the stoichiometric matrix (and the incidence matrix of the graph representing the metabolic network), $\underline{v} \in \mathbb{R}^{n_v}$ is the vector of intracellular fluxes (in mmol/gDW/h), which are assumed positive (i.e., to have a net direction), and n_s is the number of intracellular metabolites. N is assumed full-row rank, thus defining n_s independent mass balance equations. This system of equations expresses the zero balance in each internal node of the metabolic network, and imposes a set of linear equality constraints, which are not sufficient to determine a unique solution for the flux vector \underline{v} . This system of equations is often supplemented by additional mass balance equations expressing the link between the intracellular fluxes and the measurements of external fluxes (uptake or production of extracellular metabolites):

$$N_m \underline{v} = \underline{v}_m \quad (2)$$

where $\underline{v}_m \in \mathbb{R}^{n_m}$. Even though this additional information allows restricting the solution space, it is usually not sufficient to define a unique solution. More precisely, a subset of

the fluxes might be exactly calculable [2] while only intervals of values for the remaining fluxes can be computed. In general, the system of equations under consideration can be formulated as:

$$A_e \underline{v} = \underline{b}_e \quad (3)$$

$$A_i \underline{v} \leq \underline{b}_i \quad (4)$$

where the equality constraints (1) and (2) are put together in (3), and Equation (4) contains the positivity constraints as well as other bound constraints, e.g., upper bounds on some of the fluxes, corresponding to prior knowledge or biological assumptions. The matrices $A_e \in \mathbb{R}^{n_e \times n_v}$ and $A_i \in \mathbb{R}^{n_i \times n_v}$, correspond to n_e equality constraints and n_i inequality constraints, respectively. To tackle the underdeterminacy, several approaches have been proposed in recent years, which can be grouped into two distinct strategies:

1. Dealing with the underdeterminacy—this strategy is adopted in several methods where minimal and maximal bounds on the admissible fluxes are determined. This category of methods includes Flux Pathway Analysis (FPA), where convex analysis is used to decompose the admissible flux distributions into Elementary Flux Modes (EFMs) or Extreme Pathways [3,4], Flux Variability Analysis (FVA), which is a Linear-Programming (LP)-based method determining the range of admissible fluxes [5], Flux Spectrum Approach (FSA), which is another LP-based method taking insufficient and uncertain measurements into account [6]. Random sampling of the admissible solution set allows determining the marginal probability density functions of the fluxes [7–10], and statistical methods based on the maximum entropy principle can be used to infer intracellular flux distributions [11,12].
2. Reducing or eliminating the underdeterminacy—this strategy consists in adding constraints in various ways, e.g., including more measurements of the extracellular fluxes or, possibly, measurements of the intracellular fluxes using specific techniques such as ^{13}C tracing [13,14] and parallel labeling [15], leading to the sophisticated procedures of ^{13}C MFA. Alternatively, additional constraints can be introduced by formulating biological assumptions either based on prior knowledge and/or experimental observations [16,17] or systematic procedures to determine active constraints [18]. The use of thermodynamic constraints can be important in relation with reaction reversibility and the limitation of the solution space [19]. Moreover, thermodynamical constraints can prevent infeasible loops in a metabolic network as demonstrated in [20]. Underdeterminacy can also be reduced (or even eliminated) through the formulation of an optimization problem originating from the assumption of an optimal metabolic behavior of the cells. This approach corresponds to Flux Balance Analysis (FBA) [21,22], which uses an objective function expressed as a linear combination of selected fluxes. Recently, the increasing availability of metabolite profiling data obtained through gas and liquid chromatography combined with mass spectroscopy has also allowed the integration of time-course absolute quantitative metabolomics in unsteady-state (or dynamic) FBA [23,24]. In the usual situation where FBA still leads to an underdetermined system with an infinite number of flux distributions that optimize the cost function, variants of FBA have been proposed in order to define a unique solution, e.g., the geometric approach developed in [25] that searches for the minimal flux distribution satisfying the given objective. Assuming that fluxes correlate with enzyme levels, this specific flux distribution would correspond to the minimization of the amount of enzymes required to satisfy the objective defined in FBA. Ultimately, the concept of Most Accurate Fluxes [26] allows computing a unique flux distribution, hence eliminating the system underdeterminacy, with a very low computational load and without any assumption regarding an optimal biological behavior.

In the following, we review some of these methods and their implementation with a toy example, which provides a numerical illustration of the main concepts. The Matlab code of this example is provided in the Supplementary Materials associated to this article.

2. A Toy Example

Despite its small size, the metabolic network (see Figure 1) that is considered to illustrate the several methods introduced in the previous section presents many representative features, e.g., several intracellular metabolites, extracellular substrates, and intra- and extra-cellular products. This network is described by the following reactions:

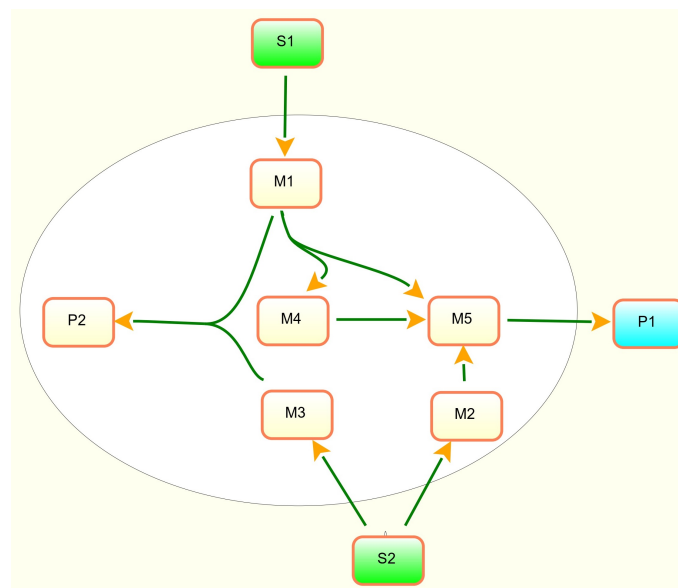
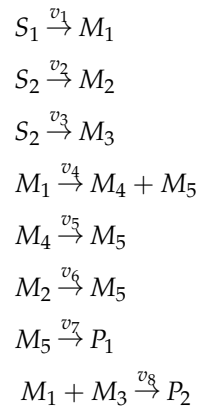


Figure 1. Simple metabolic network. M_i , $i = 1, \dots, 5$ are the internal metabolites, S_i , $i = 1, 2$ are the extracellular substrates, and P_i , $i = 1, 2$ are the extracellular and intracellular product, respectively.

The quasi-steady state assumption for the internal metabolites M_i , $i = 1, \dots, 5$, yields a system of algebraic mass-balance equations in the form of Equation (1)

$$\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{bmatrix} = N\underline{v} = 0 \quad v_i \geq 0 \quad \forall i \quad (5)$$

The stoichiometric matrix N is full row rank (i.e., 5) and the system of equations has 3 degrees of freedom (5 independent equations for 8 unknown fluxes). In an ideal measurement configuration, we consider that the 3 extracellular fluxes can be measured, for instance:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \underline{v} = N_m \underline{v} = \begin{bmatrix} v_{m1} \\ v_{m2} \\ v_{m3} \end{bmatrix} = \begin{bmatrix} 3.5 \\ 2.7 \\ 1.8 \end{bmatrix} = \underline{v}_m \quad (6)$$

so that a unique solution, i.e., flux distribution, can be found:

$$\underline{v} = \begin{bmatrix} 3.5000 \\ 0.0667 \\ 2.6333 \\ 0.8667 \\ 0.8667 \\ 0.0667 \\ 1.8000 \\ 2.6333 \end{bmatrix} \quad (7)$$

In the sequel, we will consider various situations with less measurements, and alternative methods to deal with, reduce or eliminate the underdeterminacy.

3. Dealing with the Underdeterminacy

The solution space of Equation (5) can be described using the concept of EFMs [3,27] which represent minimal, non-decomposable pathways connecting substrates to products. Every flux in the metabolic network can be described as a convex combination of EFMs:

$$\underline{v} = \sum_{i=1}^{n_{EFM}} \mu_i \underline{e}_i = E \underline{\mu}, \quad \mu_i \geq 0 \quad (8)$$

where n_{EFM} is the number of EFMs \underline{e}_i .

The EFMs can be computed using readily available software such as Metatool [28], EFMtool [29] or FluxModeCalculator [30]. The main issue associated to this computation is the combinatorial explosion of the number of EFMs with the network size (a network of less than 100 reactions can have tens of thousands of EFMs), and the fact that the computation involves some form of enumeration, which requires large computer memory space and computation time. Alternative procedures have therefore been proposed to compute minimal sets of EFMs without enumerating all of them [31]. For the small network under consideration in this review, this computation is trivial and leads to

$$E = [\underline{e}_1 \quad \underline{e}_2 \quad \underline{e}_3] = \begin{bmatrix} 0 & 1 & 0.5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.5 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (9)$$

The three EFMs define a polyhedral cone in the positive orthant, which contains all possible flux distributions. They correspond to a minimal bioreaction system, which provides an input–output representation of the cell metabolism (this kind of representation

can be very useful to derive reduced-order macroscopic representations of the culture system [32–34], but this subject will be addressed later on in this paper; see Section 5.4).



We already know that if three measurements are available, such as the ones defined in Equation (2), the solution is unique and the coefficient vector $[\mu_1 \ \mu_2 \ \mu_3]^T = [0.0667 \ 2.6333 \ 1.7333]^T$. If less measurements are available, for instance only v_{m1} and v_{m2} , then the EFM basis of the linear system

$$\begin{bmatrix} N & 0 \\ N_m & -v_m \end{bmatrix} \begin{bmatrix} v \\ 1 \end{bmatrix} = 0 \quad (11)$$

leads to the so-called extreme rays f_i [35]

$$\begin{bmatrix} f_1 & f_2 \end{bmatrix} = \begin{bmatrix} 3.5 & 3.5 \\ 0 & 2.7 \\ 2.7 & 0 \\ 0.8 & 3.5 \\ 0.8 & 3.5 \\ 0 & 2.7 \\ 1.6 & 9.7 \\ 2.7 & 0 \end{bmatrix} \quad (12)$$

which defines a pointed polyhedral cone that is a subspace of the previous solution cone. Moreover, the flux spectrum $F_0 = \{v : v_i^{\min} \leq v_i \leq v_i^{\max}\}$ with $v_i^{\min} = \min\{f_{ki}, k = 1, \dots, p\}$ and $v_i^{\max} = \max\{f_{ki}, k = 1, \dots, p\}$ is easily deduced, giving

$$\begin{bmatrix} v_1 = 3.5 \\ 0 \leq v_2 \leq 2.7 \\ 0 \leq v_3 \leq 2.7 \\ 0.8 \leq v_4 \leq 3.5 \\ 0.8 \leq v_5 \leq 3.5 \\ 0 \leq v_6 \leq 2.7 \\ 1.6 \leq v_7 \leq 9.7 \\ 0 \leq v_8 \leq 2.7 \end{bmatrix} \quad (13)$$

which indeed encloses the unique solution found with an additional measurement.

An alternative and straightforward way to compute the flux spectrum is provided by Flux Variability Analysis (FVA) [5], which consists in formulating a double optimization problem (minimization/maximization) of the flux distribution under the constraints provided by the metabolic network stoichiometry, the measurements, and any other additional biological constraints.

$$\begin{aligned} &\left. \begin{aligned} v_i^{\min} &= \min_v v_i \\ v_i^{\max} &= \max_v v_i \end{aligned} \right\} \forall i \in [1, n_v] \\ &s.t. \quad \begin{cases} A_e v = b_e \\ A_i v \leq b_i \end{cases} \end{aligned} \quad (14)$$

The unique solution to this problem, if bounded and feasible, can be computed using linear programming, as available in many software libraries (e.g., *linprog* in *Matlab* or other LP solvers such as CPLEX interfaced in the language of your choice such as Python or Julia) or dedicated environments (e.g., COBRA [36] and CellNetAnalyzer [37]). In our application

example, we use *linprog* to compute the flux spectrum in the situation where only the first measurement is available (v_{m1}) and a constraint is imposed in the form $v_2 + v_3 \leq 5$.

$$\begin{aligned} A_e &= \begin{bmatrix} & & & N & & & & & \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & b_e &= [0 \ 0 \ 0 \ 0 \ 0 \ 3.5]^T \\ A_i &= \begin{bmatrix} & & & -I_8 & & & & & \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & b_i &= [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 5]^T \end{aligned} \quad (15)$$

Thus, the spectrum F_0 is given by

$$\begin{bmatrix} v_1 = 3.5 \\ 0 \leq v_2 \leq 5 \\ 0 \leq v_3 \leq 3.5 \\ 0 \leq v_4 \leq 3.5 \\ 0 \leq v_5 \leq 3.5 \\ 0 \leq v_6 \leq 5 \\ 0 \leq v_7 \leq 12 \\ 0 \leq v_8 \leq 3.5 \end{bmatrix} \quad (16)$$

The application of FVA can be delicate when the measurements are corrupted by noise, as the constraints imposed by the metabolic network and/or the bounds on the fluxes could become incompatible with the measurement information. Several approaches take account of the measurement uncertainty, such as Flux Spectrum Analysis [6] or Adaptive Flux Variability Analysis [38], which relax the constraints to allow for a feasible solution. Here, we consider a variation of our simple example where only the first measurement $v_{m1} = 10.5$ would be available and 3 constraints would be imposed, i.e., $v_2 \leq 5$, $v_5 \leq 5$, $v_8 \leq 5$. In this case, the matrices become

$$\begin{aligned} A_e &= \begin{bmatrix} & & & N & & & & & \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & b_e &= [0 \ 0 \ 0 \ 0 \ 0 \ 10.5]^T \\ A_i &= \begin{bmatrix} & & & -I_8 & & & & & \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} & b_i &= [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 5 \ 5 \ 5]^T \end{aligned} \quad (17)$$

but unfortunately the LP solver returns the message that no feasible solution can be found. What is happening? In fact, the solution of the FVA problem without the knowledge of the first measurement v_{m1} returns the following spectrum:

$$\begin{bmatrix} 0 \leq v_1 \leq 10 \\ 0 \leq v_2 \leq 5 \\ 0 \leq v_3 \leq 5 \\ 0 \leq v_4 \leq 5 \\ 0 \leq v_5 \leq 5 \\ 0 \leq v_6 \leq 5 \\ 0 \leq v_7 \leq 15 \\ 0 \leq v_8 \leq 5 \end{bmatrix} \quad (18)$$

which shows that the maximum admissible value of v_1 is 10. Hence, the noisy measurement $v_{m1} = 10.5$ is incompatible with this upper bound, and it is not possible to include it as such in the equality constraints. A way round this issue is to introduce inequality constraints in the form

$$(1 - e)v_{m1} \leq v_1 \leq (1 + e)v_{m1} \quad (19)$$

where e represents a relative uncertainty. In our example, we could choose $e = 5\%$, which is the smallest uncertainty leading to the matrices modification

$$\begin{aligned}
 A_e &= [N] & \underline{b}_e &= [0 \ 0 \ 0 \ 0 \ 0 \ 0]^T \\
 A_i &= \begin{bmatrix} & & & -I_8 & & & & & \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & \underline{b}_i &= [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 5 \ 5 \ 5 \ -9.975 \ 11.025]^T
 \end{aligned} \tag{20}$$

and a feasible solution

$$\begin{bmatrix} 9.975 \leq v_1 \leq 10 \\ 0 \leq v_2 \leq 5 \\ 4.975 \leq v_3 \leq 5 \\ 4.975 \leq v_4 \leq 5 \\ 4.975 \leq v_5 \leq 5 \\ 0 \leq v_6 \leq 5 \\ 9.95 \leq v_7 \leq 15 \\ 4.975 \leq v_8 \leq 5 \end{bmatrix} \tag{21}$$

On another matter, it can be convenient to eliminate the equality constraints and to formulate the problem in terms of inequality constraints only in a space of reduced dimension [39]. To this end, we can use the kernel (or null space) of the matrix A_e . If $A_0 \in \mathbb{R}^{n_v \times (n_v - n_e)}$ is a matrix whose columns form a basis of this kernel ($n_q = n_v - n_e$ is the nullity of the kernel for a full row rank matrix A_e), then any solution of $A_e \underline{v} = \underline{b}_e$ (Equation (3)) can be expressed as

$$\underline{v} = \underline{v}_0 + A_0 \underline{q} \tag{22}$$

where \underline{v}_0 is a particular solution to Equation (3) and the vector $\underline{q} \in \mathbb{R}^{n_q}$ allows the reformulation of the inequality constraints as

$$A_i^0 \underline{q} \leq \underline{b}_i^0 \tag{23}$$

with $A_i^0 = A_i A_0$ and $\underline{b}_i^0 = \underline{b}_i - A_i \underline{v}_0$.

A particular solution \underline{v}_0 can for instance be obtained by solving the following problem (in this case \underline{v}_0 is the flux vector with minimum Euclidean norm in the set of solutions)

$$\begin{aligned}
 \underline{v}_0 &= \min_{\underline{v}} \underline{v}^T \underline{v} \\
 \text{s.t.} \quad & \begin{cases} A_e \underline{v} = \underline{b}_e \\ A_i \underline{v} \leq \underline{b}_i \end{cases}
 \end{aligned} \tag{24}$$

To illustrate this, we return to our original example and consider the situation where only the first measurement $v_{m1} = 3.5$ is available and a constraint is imposed in the form $v_2 + v_3 \leq 5$. \underline{v}_0 can be computed using quadratic programming, e.g., *quadprog* in Matlab

$$\underline{v}_0 = [3.500 \ 0 \ 2.625 \ 0.875 \ 0.875 \ 0 \ 1.750 \ 2.625]^T \tag{25}$$

The null space of the equality constraint matrix $A_e \in \mathbb{R}^{6 \times 8}$ can be computed using *null* in Matlab and is given by

$$A_0 = \begin{bmatrix} 0 & 0.2019 & -0.2846 & 0.2846 & 0.2846 & 0.2019 & 0.7711 & -0.2846 \\ 0 & 0.5994 & 0.2627 & -0.2627 & -0.2627 & 0.5994 & 0.0740 & 0.2627 \end{bmatrix}^T \tag{26}$$

and, in turn

$$A_i^0 = \begin{bmatrix} 0 & -0.2019 & 0.2846 & -0.2846 & -0.2846 & -0.2019 & -0.7711 & 0.2846 & -0.0827 \\ 0 & -0.5994 & -0.2627 & 0.2627 & 0.2627 & -0.5994 & -0.0740 & -0.2627 & 0.8621 \end{bmatrix}^T \quad (27)$$

and

$$\underline{b}_i^0 = [3.500 \quad 0 \quad 2.625 \quad 0.875 \quad 0.875 \quad 0 \quad 1.750 \quad 2.625 \quad 2.375]^T \quad (28)$$

The application of FVA in the q -space (a reduced space of dimension $n_q = 2$), i.e., with inequality constraints only

$$\left. \begin{aligned} q_i^{\min} &= \min_q q_i \\ q_i^{\max} &= \max_q q_i \end{aligned} \right\} \forall i \in [1, n_q] \quad (29)$$

s.t. $A_i^0 q \leq \underline{b}_i^0$

gives a spectrum Q_0

$$\begin{bmatrix} -2.3454 \leq q_1 \leq 12.9102 \\ -2.3698 \leq q_2 \leq 3.9938 \end{bmatrix} \quad (30)$$

In the reduced q -space, the system of inequality constraints defines half hyperplanes whose intersection consists of a convex polytope that contains all the admissible flux distributions \underline{q} . Uniformly sampling this convex polytope allows subsequently computing the marginal probability density functions (or marginal distributions) of each flux.

In our toy example, the rejection algorithm [40] can be applied, which boils down to uniformly sample each coordinate q_i ($\forall i \in [1, n_q]$) on $[q_i^{\min}, q_i^{\max}]$. The obtained sample \underline{q} is kept if it satisfies the inequality constraints $A_i^0 \underline{q} \leq \underline{b}_i^0$, otherwise it is rejected. The procedure is repeated until the desired number of samples is reached. Figure 2 shows 10^4 samples obtained with the rejection algorithm. Despite its simplicity and the genuine uniform distribution that it provides, this algorithm cannot be used with high dimensional spaces and irregular shaped polytopes given that the fraction of rejected samples increases dramatically with the number of metabolic fluxes considered in the network.

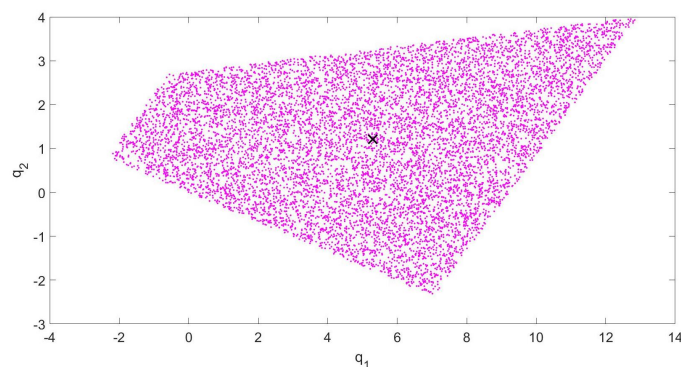


Figure 2. Rejection algorithm in the q -space (mean = X).

Other algorithms have been (and are still) developed to circumvent this problem [41,42]. Among the oldest and simplest methods, hit-and-run algorithms [43] consist of Markov Chain Monte Carlo methods that sample the convex polytope via some specific random walk. While they can be used with high dimensional spaces, their main drawback is that the samples often get stuck in some part of the polytope when this latter exhibits an irregular shape, which is generally the case. Figure 3 represents the marginal distributions of each flux in the v -space (transforming the \underline{q} samples into \underline{v} samples through Equation (22)), inferred from 10^4 samples obtained with the rejection algorithm and with a specific hit-and-run algorithm (namely, the random direction algorithm). Both results are almost identical.

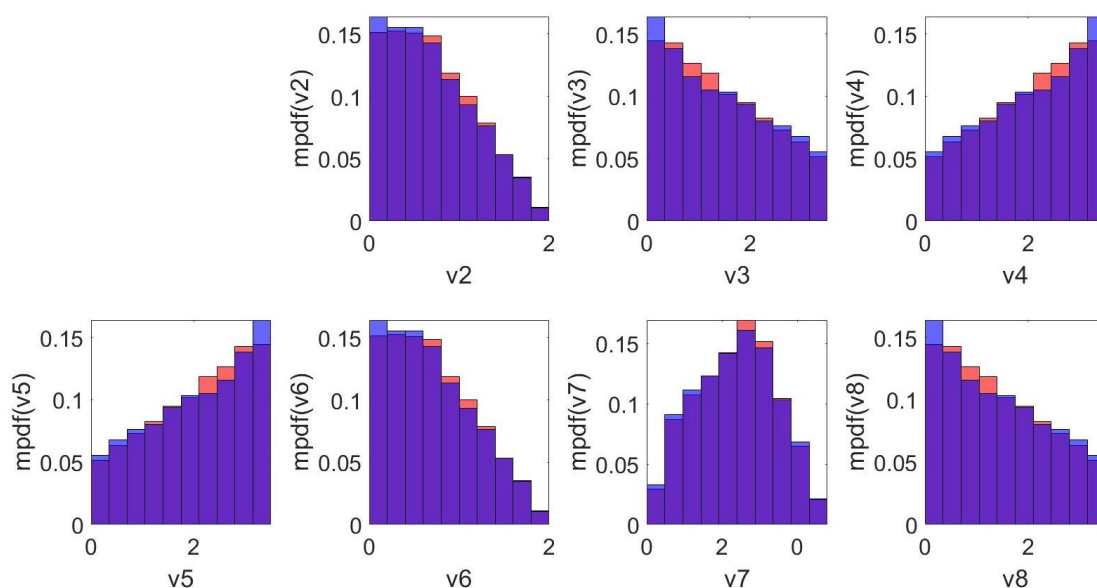


Figure 3. Marginal probability distribution in the original v -space, inferred from 10^4 samples with the rejection algorithm (red) and a hit-and-run algorithm (blue).

4. Reducing or Eliminating the Underdeterminacy

A straightforward way to reduce or eliminate underdeterminacy is to include additional measurements, either extracellular as shown in Equation (6) for our simple example, or intracellular using ^{13}C tracing [13,14] or the integration of time-course absolute quantitative metabolomics [23,24], which would amount to directly measure some of the internal fluxes v_i , $i = 1, \dots, 8$ in the toy example. However, this implies additional time-consuming, delicate, and costly experiments and equipment. If sufficient additional measurements are not available, a candidate flux distribution can be provided by Flux Balance Analysis (FBA) [21,22], which assumes some optimal behavior of the cell, such as maximum cell growth rate or maximum ATP production rate, and formulates a linear programming problem

$$\begin{aligned} \hat{v} &= \arg \min_v \underline{\lambda}^T \underline{v} \\ \text{s.t.} \quad &\begin{cases} A_e \underline{v} = \underline{b}_e \\ A_i \underline{v} \leq \underline{b}_i \end{cases} \end{aligned} \quad (31)$$

If the LP is feasible and bounded, then the cost function $J = \underline{\lambda}^T \underline{v}$ has a unique minimum value J^* , but the corresponding flux distribution \hat{v} is not necessarily unique, still implying underdeterminacy. In this latter case, it is possible to combine FBA (31) with FVA by subsequently solving a set of $2n_v$ LPs

$$\begin{aligned} \left. \begin{aligned} v_i^{\min} &= \min_{\underline{v}} v_i \\ v_i^{\max} &= \max_{\underline{v}} v_i \end{aligned} \right\} \forall i \in [1, n_v] \\ \text{s.t.} \quad &\begin{cases} A_e \underline{v} = \underline{b}_e \\ \underline{\lambda}^T \underline{v} = J^* \\ A_i \underline{v} \leq \underline{b}_i \end{cases} \end{aligned} \quad (32)$$

This FVA problem includes an additional equality constraint enforcing the value $J^* = \underline{\lambda}^T \underline{v}$ determined in the first FBA step.

This can be illustrated with our toy example, first by applying FBA with the assumption that v_7 is maximum. In this case, $\hat{v}_7 = 12$ and the corresponding flux distribution is given by

$$\hat{v} = [3.50 \quad 5.00 \quad 0 \quad 3.50 \quad 3.50 \quad 5.00 \quad 12.0 \quad 0]^T \quad (33)$$

which is confirmed to be the unique solution by applying FVA (which produces a flux spectrum which reduces to the single value \hat{v}).

If we repeat this exercise with the maximization of v_8 , we find $\hat{v}_8 = 3.5$, but the corresponding flux distribution is not unique and belongs to the spectrum

$$\begin{bmatrix} v_1 = 3.5 \\ 0 \leq v_2 \leq 1.5 \\ v_3 = 3.5 \\ v_4 = 0 \\ v_5 = 0 \\ 0 \leq v_6 \leq 1.5 \\ 0 \leq v_7 \leq 1.5 \\ v_8 = 3.5 \end{bmatrix} \quad (34)$$

5. An Overview of Important Topics

In this section, we address a few important questions when dealing with the analysis of metabolic networks. Some of them have a direct impact on the underdeterminacy of the metabolic flux analysis, such as the topology and size of the metabolic network or the selection of the extra- or intra-cellular measurements. On the contrary, other issues, such as the dynamic evolution of the extracellular fluxes, might have no particular influence on this underdeterminacy, but will impact the computational procedure and the visualization of the results, and this is why we include them in the global picture.

5.1. How to Select the Size/Detail of the Metabolic Network?

The structure and size of the metabolic network can be represented by an incidence graph, whose topological properties are important for the solution of Equations (1)–(4). The analysis of these properties, such as determinacy, redundancy, balanceability, and calculability, can be assessed using software tools such as CellNetAnalyser [37] or COPASI [44], which were the first to propose such functionalities. The size of the network will also have a tremendous influence on the number of EFMs. A large network will probably imply that the EFMs are no longer enumerable due to memory and computation limitations. Hence, the importance of procedures to compute only subsets of EFMs such as [31], and of concepts such as the giant strong component (GSC), which represents the largest fully connected part of a metabolic network as introduced originally in [45]. Indeed, the GSC usually contains less than one-third of the nodes of the network, but key metabolites, and is more feasible for analysis of flux distribution and computation of EFMs. Another concept of interest is the introduction of minimal cut sets (MCSs), which represent sets of reactions whose removal will disable certain network functions [46], and which have been shown to be the EFMs of a dual network [47]. MCSs can be used to study the observability of reaction rates in metabolic flux analyses. The selection of the network will therefore be a compromise between describing the metabolic features of interest and the tractability of the computation procedure underlying the flux determination. An open research question is the selection of the right metabolic network for the derivation of low-order dynamic models (as introduced in the following Section 5.4). Should a reduction be operated a priori based on biological assumptions and simplifications or only a posteriori, in the course of the derivation of the dynamic macroscopic model?

5.2. Dynamic Metabolic Flux Interval Analysis

To date, we have considered that the measured specific extracellular fluxes are constant. This is typically the case in the early exponential growth phase of batch cultures or in

the steady state of continuous cultures. However, there are other situations where the fluxes are time varying following changes in the environmental conditions (substrate excess or depletion, accumulation of byproducts). This can occur in fed-batch cultures, the transient phase between batch and continuous modes, or the transient phase between different setpoints in continuous operation. The previous analysis can be extended to these situations by considering a time-scale separation where the bioreactor environment is the slow subsystem and the cells or micro-organisms the fast subsystem, which can therefore be considered in a pseudo steady-state. The dynamics of the extracellular substrates S and products P can be described by a set of mass balance equations

$$\begin{aligned}\frac{d\underline{S}}{dt} &= -\underline{v}_S X - D(\underline{S} - \underline{S}_{in}) \\ \frac{d\underline{P}}{dt} &= \underline{v}_P X - D(\underline{P} - \underline{P}_{in})\end{aligned}\quad (35)$$

where \underline{S} and \underline{P} represent the extracellular substrate and product concentrations, respectively, \underline{v}_S is the vector of specific uptake rates, \underline{v}_P the vector of specific production rates, $D = F_{in}/V$ is the dilution rate (ratio of the inlet flow rate $F_{in}(t)$ to the culture volume $V(t)$).

A straightforward approach consists in smoothing the extracellular concentration evolutions, computing the derivatives of the smoothed signals and evaluating the uptake and production rates using Equation (35) and the knowledge about the transportation terms (functions of the dilution rate and inlet concentrations \underline{S}_{in} and \underline{P}_{in}). This approach was originally developed for MFA with no underdeterminacy or even overdeterminacy. One of the earlier reports can be found in [48] where the lysine fermentation process by *Corynebacterium glutamicum* is studied and the cell metabolic state is estimated online based on a small metabolic network with 11 fluxes. This work is extended in [49], where HEK cells are cultivated in perfusion and the metabolic fluxes are estimated online using a medium-size metabolic network of 40 reactions. Other notable accounts include [50], where *Escherichia coli* cultivations shifting from carbon limitation to nitrogen limitation and vice versa are studied, and [51], where a human cell line is analyzed and the authors compare Dynamic MFA (DMFA) to a flux average approach where the culture is divided in phases over which constant (average) fluxes are considered.

When underdeterminacy prevails, the approaches previously reviewed can be extended to take account of the dynamic evolution of the extracellular fluxes as well.

In [52,53], Dynamic Flux Balance Analysis (DFBA) is introduced and applied to the analysis of diauxic growth of *Escherichia coli* on glucose and acetate. Two optimization approaches can be considered: (a) a sequence of static LP problems corresponding to the subdivision of the batch time into time intervals over which the fluxes are assumed constant, or (b) a global dynamic optimization formulated over the total batch duration that can be solved using multiple shooting and orthogonal collocation to be converted into a NLP. The latter approach allows the consideration of an integrated (over the system trajectory) optimality objective or an end-of-batch objective, as well as nonlinear constraints on the fluxes resulting from a priori knowledge about kinetic expressions, but results in a significantly more complex problem. DFBA is nowadays a popular approach, which has led to many interesting applications (see for instance [54–58]) and the emergence of software tools such as DFBLAB [59].

In [60], Flux Spectrum Approach (FSA), which is a method based on the formulation of a set of min/max LP problems taking account of a range of measurement errors, is applied to the analysis of the evolution over time of the fluxes in small metabolic network of the CHO metabolism and data borrowed from [35].

In [61], a linear objective function subject to elementary modes as constraints is optimized to determine the fluxes in the metabolic network of *Corynebacterium glutamicum* at different temporal phases of fermentation. The use of convex analysis and EFMs is investigated in [35] to determine intervals for the metabolic fluxes of CHO cells in batch cultures, which are decomposed into several time periods corresponding to different phases of the

cell life cycle. This work is extended to more detailed metabolic networks in [62] and to the dynamic evolution of the flux spectrum, without assuming a decomposition in phases, in [34] where convex analysis is applied to hybridoma cultures switching from batch to perfusion mode.

5.3. How to Represent the Accumulation of Internal Metabolites?

Besides the possible time evolution of the extracellular fluxes discussed in the previous subsection, another dynamic phenomenon might need special attention: the accumulation of an intracellular metabolite over time, implying that the traditional assumption of quasi steady-state is no longer valid. This situation is, for instance, observed in cultures of photosynthetic micro-organisms that can accumulate various components, such as carbohydrates and lipids. It is also well-known that yeasts, such as *Saccharomyces cerevisiae*, are able to accumulate some carbohydrates, e.g., trehalose, that play a role of carbon and energy reserve, as well as of stress protectant against harmful environmental conditions [63]. Abandoning the quasi steady state assumption for some of the intracellular metabolites boils down to removing their corresponding rows in the stoichiometric matrix N involved in Equation (1), hence increasing the number of degrees of freedom that characterizes the system underdeterminacy. To compensate for the mass balance equations withdrawn from (1), the kinetics of accumulation and/or reuse of the intracellular metabolites should be included in mass balance ODEs. However, the lack of intracellular measurements along time usually hampers the identification of these intracellular reaction rates. In [64], the authors propose the DRUM (Dynamic Reduction of Unbalanced Metabolism) modeling framework that consists in defining subsets of balanced intracellular metabolites that are interconnected via linking metabolites. These latter may accumulate or be reused. The subsets of balanced metabolites are reduced to macroscopic reactions, based on Elementary Flux Mode analysis, for which simple kinetic models are derived that allow building mass balance ODEs for the linking metabolites, biomass production, substrate consumption and product excretion. This methodology is applied to the lipid and carbohydrate accumulation in the microalgae *Tisochrysis lutea*. The modeling approach proposed in [65] can be applied to any metabolic network whose kinetics can be locally linearized. As in [64], the authors also consider subnetworks of fast reactions (involving metabolites that are assumed at quasi steady state) connected via metabolites consumed at low rates. Based on this time scale separation, the methodology allows reducing high dimensional linearized models, while accounting for the accumulation of some metabolites, as well as for the dilution effect due to biomass growth. In [66], a FBA-based simulator of *Saccharomyces cerevisiae* fed-batch cultures is proposed, using the assumption that the intracellular alpha-ketoglutarate is unbalanced. Its dynamics are described with a linear combination of the glucose and nitrogen specific uptake rates whose models involve the alpha-ketoglutarate concentration.

5.4. Model Reduction to Macroscopic Scale

Macroscopic models mainly predict biomass growth, the consumption of external substrates and the secretion of external products. Their structural simplicity allows their use for bioprocess optimization, control and online monitoring. To these purposes, it can be worthy to reduce metabolic models to a macroscopic scale. In [67,68], the authors propose a systematic methodology to build macroscopic reaction rate models via the definition of additional constraints whose number corresponds exactly to the number of degrees of freedom, i.e., the difference between the number of metabolic fluxes and the number of available equality constraints, hence removing the system underdeterminacy. These constraints express linear combinations of the metabolic fluxes as nonlinear functions of the extracellular concentrations, which correspond to the macroscopic reaction rate models. Another method that has often been used to define macroscopic reactions is the Elementary Flux Mode analysis [32–34,69]. As discussed above, EFMs are the shortest pathways from substrates to products. However, the main drawback with EFMs is that

their total number dramatically increases with the network size. To overcome this problem, different solutions have been proposed. To avoid the exhaustive enumeration of all EFMs, refs [31,70] propose fast algorithms that randomly compute minimal sets of EFMs. In [71], the number of EFMs is reduced through a projection from the flux space to the yield space. Refs. [72,73] select EFMs, via ranking or controlled random search algorithms, using a multi-criteria objective function that combines prediction error, model size and efficiency of the EFMs (investment required to produce the enzymes). Another approach is made of Lumped Hybrid Cybernetic Models in which EFMs are grouped into clusters, each of them being associated to an average EFM [74]. In [75], the column generation method is used to determine a (non-necessarily unique) minimal subset of EFMs, which consists in solving iteratively two levels of optimization problems: the master problem is a quadratic optimization problem that identifies the macroscopic flux values using a subset of EFMs, and the subproblem is a linear problem for identifying EFMs that improve the model prediction in the master problem. Based on extensions of Dynamic Metabolic Flux Analysis introduced in [76], that only uses concentration measurements and avoids any numerical differentiation, refs. [77,78] select reduced sets of EFMs via a geometrical reduction (excluding EFMs with a cosine-similarity algorithm) followed by a multi-objective genetic algorithm that minimizes the prediction error and the size of the EFMs subset. A linear optimization problem has been formulated in [79] for selecting the best subset of EFMs based on a relaxation criterion. The methodology is extended in [80] and includes a more efficient selection procedure for the minimal subset of EFMs. Note that once the macroscopic reactions have been deduced from the metabolic network, it remains to identify their kinetic models. To that purpose, general kinetic models and systematic identification procedures can be very useful [81,82]. Model reduction methodologies based on subsets of balanced metabolites interconnected via linking metabolites that may accumulate within cells [64,65], have also been introduced in the previous section. Finally, independently of any EFM computation, macroscopic models may also consist of ODEs describing the mass balances for the biomass and the extracellular species, in which the reaction rates are computed at each time point by solving a FBA problem [16,17,66]. In [66], the underdeterminacy of the FBA problem is taken into account within the set of mass balance ODEs by computing the minimum and maximum admissible values of the ethanol concentration associated to the respective minimum and maximum admissible values of the specific ethanol production rate that are computed through FVA. Hence, the underdeterminacy at the level of FBA leads to corridors of admissible values along time for the concentrations of some macroscopic components, typically the extracellular products that are secreted.

5.5. How to Handle the Measurement Errors?

In the toy example, the adverse effect of an error on the measured extracellular flux was alleviated by relaxing an equality constraint and reformulating it as two inequality constraints (see Equation (19)). Such errors are frequent in practice, as it is necessary to compute the specific extracellular fluxes from the measurements of the evolution of the biomass and the extracellular concentrations. If we consider, for simplicity, the exponential growth phase of a batch culture, this can be formulated in the following way:

$$\begin{aligned}\frac{dS}{dt} &= -v_S X \\ \frac{dP}{dt} &= v_P X \\ \frac{dX}{dt} &= \mu X\end{aligned}\tag{36}$$

The solution of these equations are in the form of a linear regression

$$\begin{aligned}\underline{S}(t) &= -\frac{v_S}{\mu} X(t) + (\underline{S}(0) + \frac{v_S}{\mu} X(0)) = a_1 X(t) + b_1 \\ \underline{P}(t) &= \frac{v_P}{\mu} X(t) + (\underline{P}(0) - \frac{v_P}{\mu} X(0)) = a_2 X(t) + b_2 \\ \ln(X(t)) &= \mu t + \ln(X(0)) = a_3 t + b_3\end{aligned}\quad (37)$$

where the regressors a_i give the estimation of the specific fluxes. If the environmental conditions evolve over time, it will be necessary to consider the inflows and outflows of the bioreactor as in dynamic model (35), and the variation of the growth rate of the biomass. This will imply the evaluation of the time derivatives using smoothing and numerical differentiation as explained in Section 5.2, or the formulation of the fluxes as piecewise linear functions without the need for numerical differentiation as proposed by [76]. Whatever the numerical procedure, experimental and numerical errors will always corrupt the information about the specific fluxes, which could in turn become incompatible with the constraints imposed by the metabolic network and prior knowledge about the system biology. This has to be taken into account by some form of constraints relaxation such as developed in Flux Spectrum Analysis [6] or Adaptive Flux Variability Analysis [38], where the uncertainty on the extracellular fluxes is represented as an interval around the measured values. In [38], minimum values for these uncertainties are determined by solving a sequence of optimization problems. The impact of these uncertainties on the solution, i.e., on the extent of the intervals for the intracellular fluxes, will largely depend on the structure and size of the metabolic network. This point is, however, still an open research question.

5.6. Some Further Perspectives on Sampling Algorithms

The ongoing research on sampling algorithms aims at improving their computational efficiency, convergence properties and ability to extensively explore the convex polytope of admissible flux distributions. To avoid the above mentioned problem of the samples that often become stuck in some part of the polytope when using hit-and-run algorithms [43], other methods have been proposed such as the artificial centering hit-and-run method (ACHR) [83] and the optimized general parallel sampler (OPTGP) [9]. Ref. [84] showed that the ACHR algorithms have convergence problems with high-dimensional polytopes, and introduced rounding methods with better performances. These methods were further improved in [10], leading to the coordinated hit-and-run with rounding (CHRR) method that computes the largest ellipsoid inscribed in the polytope and the rounding transformation of this ellipsoid into a unit ball. This latter transformation is then applied to the convex polytope whose sampling therefore becomes much more efficient in terms of computational time and convergence. Regarding these criteria, refs. [41,42] have shown that CHRR outperforms ACHR and OPTGP. Recently, refs. [39,85] have proposed the DISCOPOLIS (DIScrete Sampling of CONVex POLYtopes via Linear program Iterative Sequences) algorithm that, instead of being a Markov Chain Monte Carlo method, provides (subsets of) samples that are independent of each other. This allows obtaining larger ranges of admissible flux values. Given the increasing complexity of the available metabolic networks, involving thousands of fluxes, there is still a need for the development and improvement of sampling algorithms with a reasonably low computational time and extensive exploration abilities for irregular shaped polytopes. Another difficult challenge consists in sampling non-convex solution spaces that are observed when using thermodynamical constraints for preventing infeasible loops in the metabolic network [20,84].

6. Conclusions

This paper reviews and applies to a toy example (the Matlab code of this example is provided in the Supplementary Materials associated to this article) some methods that can be used with underdetermined problems in metabolic flux analysis. These methods

are grouped in two different strategies, the former consisting in simply dealing with the underdeterminacy without trying to reduce it, and the latter consisting in reducing or even eliminating the underdeterminacy. The choice between these two strategies depends on the specific problem to be solved and on the goals of the analysis. On the one hand, one could be interested by a unique solution that would be representative of a specific metabolic behavior in the considered cell line, e.g., by adding additional constraints describing these specific conditions and/or by defining an appropriate optimal behavior through FBA. On the other hand, one could be interested by the diversity of the possible metabolic behaviors resulting from the underdeterminacy, e.g., by analyzing the marginal distributions of each flux obtained from a sampling method. The point of utmost importance is to be aware of the system underdeterminacy. For example, even if one solution is provided by an algorithm used to solve the linear program of a FBA problem, that solution is not necessarily unique and a quick check via FVA could highlight that the system remains underdetermined. Many of the methods that have been presented are actually complementary, as illustrated in the abovementioned coupling of FBA with FVA. System underdeterminacy is a direct consequence of biological complexity and of the limited access to intracellular measurements. Developing methods to analyze underdetermined systems and/or to reduce their underdeterminacy in diverse and complementary ways will therefore remain an important research topic in the future. The interested reader might also consider the recent review [86], which includes an in-depth discussion of kinetic approaches for modeling cell metabolism.

Supplementary Materials: The following material is available online at <https://www.mdpi.com/article/10.3390/pr9091577/s1>, code S1: toy_example.zip.

Author Contributions: Conceptualization, P.B. and A.V.W.; methodology, P.B. and A.V.W.; software, P.B. and A.V.W.; investigation, P.B. and A.V.W.; writing—original draft preparation, P.B. and A.V.W.; writing—review and editing, P.B. and A.V.W.; visualization, P.B. and A.V.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stephanopoulos, G.; Aristidou, A.A.; Nielsen, J. *Metabolic Engineering: Principles and Methodologies*; Elsevier Science: San Diego, CA, USA, 1998.
2. Stelling, J.; Klamt, S.; Bettenbrock, K.; Schuster, S.; Gilles, E.D. Metabolic network structure determines key aspects of functionality and regulation. *Nature* **2002**, *420*, 190–193. [[CrossRef](#)]
3. Schuster, S.; Hilgetag, C. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* **1994**, *2*, 165–182. [[CrossRef](#)]
4. Klamt, S.; Stelling, J. Two approaches for metabolic pathway analysis? *Trends Biotechnol.* **2003**, *21*, 64–69. [[CrossRef](#)]
5. Mahadevan, R.; Schilling, C. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **2003**, *5*, 264–276. [[CrossRef](#)] [[PubMed](#)]
6. Llaneras, F.; Picó, J. An interval approach for dealing with flux distributions and elementary modes activity patterns. *J. Theor. Biol.* **2007**, *246*, 290–308. [[CrossRef](#)] [[PubMed](#)]
7. Thiele, I.; Price, N.D.; Vo, T.D.; Palsson, B.Ø. Candidate metabolic network states in human mitochondria impact of diabetes, ischemia, and diet. *J. Biol. Chem.* **2005**, *280*, 11683–11695. [[CrossRef](#)]
8. Saa, P.A.; Nielsen, L.K. ll-ACHRB: A scalable algorithm for sampling the feasible solution space of metabolic networks. *Bioinformatics* **2016**, *32*, 2330–2337. [[CrossRef](#)] [[PubMed](#)]
9. Megchelenbrink, W.; Huynen, M.; Marchiori, E. optGpSampler: An improved tool for uniformly sampling the solution-space of genome-scale metabolic networks. *PLoS ONE* **2014**, *9*, e86587. [[CrossRef](#)] [[PubMed](#)]
10. Haraldsdóttir, H.S.; Cousins, B.; Thiele, I.; Fleming, R.M.; Vempala, S. CHRR: Coordinate hit-and-run with rounding for uniform sampling of constraint-based models. *Bioinformatics* **2017**, *33*, 1741–1743. [[CrossRef](#)] [[PubMed](#)]
11. De Martino, A.; De Martino, D. An introduction to the maximum entropy approach and its application to inference problems in biology. *Heliyon* **2018**, *4*, e00596. [[CrossRef](#)]
12. De Martino, D.; Andersson, A.M.; Bergmiller, T.; Guet, C.C.; Tkačik, G. Statistical mechanics for metabolic networks during steady state growth. *Nat. Commun.* **2018**, *9*, 2988. [[CrossRef](#)] [[PubMed](#)]

13. Ahn, W.S.; Antoniewicz, M.R. Metabolic flux analysis of CHO cells at growth and non-growth phases using isotopic tracers and mass spectrometry. *Metab. Eng.* **2011**, *13*, 598–609. [[CrossRef](#)] [[PubMed](#)]
14. Long, C.P.; Antoniewicz, M.R. High-resolution ¹³C metabolic flux analysis. *Nat. Protoc.* **2019**, *14*, 2856–2877. [[CrossRef](#)]
15. Crown, S.B.; Antoniewicz, M.R. Parallel labeling experiments and metabolic flux analysis: Past, present and future methodologies. *Metab. Eng.* **2013**, *16*, 21–32. [[CrossRef](#)]
16. Richelle, A.; Gziri, K.M.; Bogaerts, P. A methodology for building a macroscopic FBA-based dynamical simulator of cell cultures through flux variability analysis. *Biochem. Eng. J.* **2016**, *114*, 50–64. [[CrossRef](#)]
17. Bogaerts, P.; Gziri, K.M.; Richelle, A. From MFA to FBA: Defining linear constraints accounting for overflow metabolism in a macroscopic FBA-based dynamical model of cell cultures in bioreactor. *J. Process Control* **2017**, *60*, 34–47. [[CrossRef](#)]
18. Nikdel, A.; Braatz, R.D.; Budman, H.M. A systematic approach for finding the objective function and active constraints for dynamic flux balance analysis. *Bioprocess Biosyst. Eng.* **2018**, *41*, 641–655. [[CrossRef](#)] [[PubMed](#)]
19. Soh, K.C.; Hatzimanikatis, V. Constraining the flux space using thermodynamics and integration of metabolomics data. In *Metabolic Flux Analysis*; Springer Science+Business Media: New York, NY, USA, 2014; pp. 49–63.
20. De Martino, D. Thermodynamics of biochemical networks and duality theorems. *Phys. Rev. E* **2013**, *87*, 052108. [[CrossRef](#)]
21. Raman, K.; Chandra, N. Flux balance analysis of biological systems: Applications and challenges. *Br. Bioinform.* **2009**, *10*, 435–449. [[CrossRef](#)]
22. Orth, J.D.; Thiele, I.; Palsson, B.Ø. What is flux balance analysis? *Nat. Biotechnol.* **2010**, *28*, 245–248. [[CrossRef](#)]
23. Willemsen, A.M.; Hendrickx, D.M.; Hoefsloot, H.C.; Hendriks, M.M.; Wahl, S.A.; Teusink, B.; Smilde, A.K.; van Kampen, A.H. MetDFBA: Incorporating time-resolved metabolomics measurements into dynamic flux balance analysis. *Mol. BioSyst.* **2015**, *11*, 137–145. [[CrossRef](#)] [[PubMed](#)]
24. Bordbar, A.; Yurkovich, J.T.; Paglia, G.; Rolfsson, O.; Sigurjónsson, Ó.E.; Palsson, B.O. Elucidating dynamic metabolic physiology through network integration of quantitative time-course metabolomics. *Sci. Rep.* **2017**, *7*, 46249. [[CrossRef](#)] [[PubMed](#)]
25. Smallbone, K.; Simeonidis, E. Flux balance analysis: A geometric perspective. *J. Theor. Biol.* **2009**, *258*, 311–315. [[CrossRef](#)] [[PubMed](#)]
26. Gziri, K.M.; Bogaerts, P. Determining a unique solution to underdetermined metabolic networks via a systematic path through the Most Accurate Fluxes. *IFAC-PapersOnLine* **2019**, *52*, 352–357. [[CrossRef](#)]
27. Zanghellini, J.; Ruckerbauer, D.E.; Hanscho, M.; Jungreuthmayer, C. Elementary flux modes in a nutshell: Properties, calculation and applications. *Biotechnol. J.* **2013**, *8*, 1009–1016. [[CrossRef](#)]
28. Pfeiffer, T.; Sanchez-Valdenebro, I.; Nuno, J.; Montero, F.; Schuster, S. METATOOL: For studying metabolic networks. *Bioinformatics* **1999**, *15*, 251–257. [[CrossRef](#)]
29. Terzer, M.; Stelling, J. Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics* **2008**, *24*, 2229–2235. [[CrossRef](#)] [[PubMed](#)]
30. Van Klinken, J.B.; Willems van Dijk, K. FluxModeCalculator: An efficient tool for large-scale flux mode computation. *Bioinformatics* **2016**, *32*, 1265–1266. [[CrossRef](#)]
31. Jungers, R.M.; Zamorano, F.; Blondel, V.D.; Vande Wouwer, A.; Bastin, G. Fast computation of minimal elementary decompositions of metabolic flux vectors. *Automatica* **2011**, *47*, 1255–1259. [[CrossRef](#)]
32. Provost, A.; Bastin, G. Dynamic metabolic modelling under the balanced growth condition. *J. Process Control* **2004**, *14*, 717–728. [[CrossRef](#)]
33. Zamorano, F.; Vande Wouwer, A.; Jungers, R.M.; Bastin, G. Dynamic metabolic models of CHO cell cultures through minimal sets of elementary flux modes. *J. Biotechnol.* **2013**, *164*, 409–422. [[CrossRef](#)] [[PubMed](#)]
34. Fernandes de Sousa, S.; Bastin, G.; Jolicoeur, M.; Vande Wouwer, A. Dynamic metabolic flux analysis using a convex analysis approach: Application to hybridoma cell cultures in perfusion. *Biotechnol. Bioeng.* **2016**, *113*, 1102–1112. [[CrossRef](#)] [[PubMed](#)]
35. Provost, A. *Metabolic Design of Dynamic Bioreaction Models*; Faculté des Sciences Appliquées, Université catholique de Louvain: Louvain-la-Neuve, Belgium, 2006.
36. Becker, S.A.; Feist, A.M.; Mo, M.L.; Hannum, G.; Palsson, B.Ø.; Herrgard, M.J. Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox. *Nat. Protoc.* **2007**, *2*, 727–738. [[CrossRef](#)] [[PubMed](#)]
37. Klamt, S.; Saez-Rodriguez, J.; Gilles, E.D. Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Syst. Biol.* **2007**, *1*, 1–13. [[CrossRef](#)]
38. Abbate, T.; Dewasme, L.; Vande Wouwer, A.; Bogaerts, P. Adaptive flux variability analysis of HEK cell cultures. *Comput. Chem. Eng.* **2020**, *133*, 106633. [[CrossRef](#)]
39. Bogaerts, P.; Rooman, M. DISCOPOLIS: An algorithm for uniform sampling of metabolic flux distributions via iterative sequences of linear programs. *IFAC-PapersOnLine* **2019**, *52*, 269–274. [[CrossRef](#)]
40. Rubinstein, R. Generating random vectors uniformly distributed inside and on the surface of different regions. *Eur. J. Oper. Res.* **1982**, *10*, 205–209. [[CrossRef](#)]
41. Herrmann, H.A.; Dyson, B.C.; Vass, L.; Johnson, G.N.; Schwartz, J.M. Flux sampling is a powerful tool to study metabolism under changing environmental conditions. *NPJ Syst. Biol. Appl.* **2019**, *5*, 1–8. [[CrossRef](#)]
42. Fallahi, S.; Skaug, H.J.; Alendal, G. A comparison of Monte Carlo sampling methods for metabolic network models. *PLoS ONE* **2020**, *15*, e0235393. [[CrossRef](#)]

43. Smith, R.L. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper. Res.* **1984**, *32*, 1296–1308. [[CrossRef](#)]
44. Hoops, S.; Sahle, S.; Gauges, R.; Lee, C.; Pahle, J.; Simus, N.; Singhal, M.; Xu, L.; Mendes, P.; Kummer, U. COPASI—A complex pathway simulator. *Bioinformatics* **2006**, *22*, 3067–3074. [[CrossRef](#)]
45. Ma, H.W.; Zeng, A.P. The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* **2003**, *19*, 1423–1430. [[CrossRef](#)] [[PubMed](#)]
46. Klamt, S.; Gilles, E.D. Minimal cut sets in biochemical reaction networks. *Bioinformatics* **2004**, *20*, 226–234. [[CrossRef](#)] [[PubMed](#)]
47. Ballerstein, K.; von Kamp, A.; Klamt, S.; Haus, U.U. Minimal cut sets in a metabolic network are elementary modes in a dual network. *Bioinformatics* **2012**, *28*, 381–387. [[CrossRef](#)]
48. Takiguchi, N.; Shimizu, H.; Shioya, S. An on-line physiological state recognition system for the lysine fermentation process based on a metabolic reaction model. *Biotechnol. Bioeng.* **1997**, *55*, 170–181. [[CrossRef](#)]
49. Henry, O.; Kamen, A.; Perrier, M. Monitoring the physiological state of mammalian cell perfusion processes by on-line estimation of intracellular fluxes. *J. Process Control* **2007**, *17*, 241–251. [[CrossRef](#)]
50. Lequeux, G.; Beauprez, J.; Maertens, J.; Van Horen, E.; Soetaert, W.; Vandamme, E.; Vanrolleghem, P.A. Dynamic metabolic flux analysis demonstrated on cultures where the limiting substrate is changed from carbon to nitrogen and vice versa. *J. Biomed. Biotechnol.* **2010**, *2010*, 621645. [[CrossRef](#)] [[PubMed](#)]
51. Niklas, J.; Schröder, E.; Sandig, V.; Noll, T.; Heinzle, E. Quantitative characterization of metabolism and metabolic shifts during growth of the new human cell line AGE1. HN using time resolved metabolic flux analysis. *Bioprocess Biosyst. Eng.* **2011**, *34*, 533–545. [[CrossRef](#)]
52. Varma, A.; Palsson, B.O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **1994**, *60*, 3724–3731. [[CrossRef](#)]
53. Mahadevan, R.; Edwards, J.S.; Doyle III, F.J. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* **2002**, *83*, 1331–1340. [[CrossRef](#)]
54. Hjersted, J.L.; Henson, M.A. Optimization of fed-batch *Saccharomyces cerevisiae* fermentation using dynamic flux balance models. *Biotechnol. Prog.* **2006**, *22*, 1239–1248. [[CrossRef](#)] [[PubMed](#)]
55. Meadows, A.L.; Karnik, R.; Lam, H.; Forestell, S.; Snedecor, B. Application of dynamic flux balance analysis to an industrial *Escherichia coli* fermentation. *Metab. Eng.* **2010**, *12*, 150–160. [[CrossRef](#)]
56. Hanly, T.J.; Henson, M.A. Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. *Biotechnol. Bioeng.* **2011**, *108*, 376–385. [[CrossRef](#)]
57. Grafahrend-Belau, E.; Junker, A.; Eschenröder, A.; Müller, J.; Schreiber, F.; Junker, B.H. Multiscale metabolic modeling: Dynamic flux balance analysis on a whole-plant scale. *Plant Physiol.* **2013**, *163*, 637–647. [[CrossRef](#)]
58. Emenike, V.N.; Schenkendorf, R.; Krewer, U. Model-based optimization of biopharmaceutical manufacturing in *Pichia pastoris* based on dynamic flux balance analysis. *Comput. Chem. Eng.* **2018**, *118*, 1–13. [[CrossRef](#)]
59. Gomez, J.A.; Höffner, K.; Barton, P.I. DFBalab: A fast and reliable MATLAB code for dynamic flux balance analysis. *BMC Bioinform.* **2014**, *15*, 1–10. [[CrossRef](#)]
60. Llaneras, F.; Picó, J. A procedure for the estimation over time of metabolic fluxes in scenarios where measurements are uncertain and/or insufficient. *BMC Bioinform.* **2007**, *8*, 1–25. [[CrossRef](#)]
61. Gayen, K.; Venkatesh, K.V. Analysis of optimal phenotypic space using elementary modes as applied to *Corynebacterium glutamicum*. *BMC Bioinform.* **2006**, *7*, 1–13. [[CrossRef](#)]
62. Zamorano, F.; Vande Wouwer, A.; Bastin, G. A detailed metabolic flux analysis of an underdetermined network of CHO cells. *J. Biotechnol.* **2010**, *150*, 497–508. [[CrossRef](#)]
63. Richelle, A.; Fickers, P.; Bogaerts, P. Macroscopic modelling of baker's yeast production in fed-batch cultures and its link with trehalose production. *Comput. Chem. Eng.* **2014**, *61*, 220–233. [[CrossRef](#)]
64. Baroukh, C.; Muñoz-Tamayo, R.; Steyer, J.P.; Bernard, O. DRUM: A new framework for metabolic modeling under non-balanced growth. Application to the carbon metabolism of unicellular microalgae. *PLoS ONE* **2014**, *9*, e104499.
65. López Zazueta, C.; Bernard, O.; Gouzé, J.L. Dynamical reduction of linearized metabolic networks through quasi steady state approximation. *AIChE J.* **2019**, *65*, 18–31. [[CrossRef](#)]
66. Plaza, J.; Bogaerts, P. FBA-based simulator of *Saccharomyces cerevisiae* fed-batch cultures involving an internal unbalanced metabolite. *IFAC-PapersOnLine* **2019**, *52*, 169–174. [[CrossRef](#)]
67. Haag, J.E.; Vande Wouwer, A.; Bogaerts, P. Systematic procedure for the reduction of complex biological reaction pathways and the generation of macroscopic equivalents. *Chem. Eng. Sci.* **2005**, *60*, 459–465. [[CrossRef](#)]
68. Haag, J.E.; Vande Wouwer, A.; Bogaerts, P. Dynamic modeling of complex biological systems: A link between metabolic and macroscopic description. *Math. Biosci.* **2005**, *193*, 25–49. [[CrossRef](#)]
69. Niu, H.; Amribt, Z.; Fickers, P.; Tan, W.; Bogaerts, P. Metabolic pathway analysis and reduction for mammalian cell cultures—Towards macroscopic modeling. *Chem. Eng. Sci.* **2013**, *102*, 461–473. [[CrossRef](#)]
70. Machado, D.; Soons, Z.; Patil, K.R.; Ferreira, E.C.; Rocha, I. Random sampling of elementary flux modes in large-scale metabolic networks. *Bioinformatics* **2012**, *28*, i515–i521. [[CrossRef](#)]
71. Song, H.S.; Ramkrishna, D. Reduction of a set of elementary modes using yield analysis. *Biotechnol. Bioeng.* **2009**, *102*, 554–568. [[CrossRef](#)] [[PubMed](#)]

72. Soons, Z.I.; Ferreira, E.C.; Rocha, I. Selection of elementary modes for bioprocess control. *IFAC Proc. Vol.* **2010**, *43*, 156–161. [[CrossRef](#)]
73. Soons, Z.I.; Ferreira, E.C.; Rocha, I. Identification of minimal metabolic pathway models consistent with phenotypic data. *J. Process Control* **2011**, *21*, 1483–1492. [[CrossRef](#)]
74. Song, H.S.; Ramkrishna, D.; Pinchuk, G.E.; Beliaev, A.S.; Konopka, A.E.; Fredrickson, J.K. Dynamic modeling of aerobic growth of *Shewanella oneidensis*. Predicting triaouxic growth, flux distributions, and energy requirement for growth. *Metab. Eng.* **2013**, *15*, 25–33. [[CrossRef](#)] [[PubMed](#)]
75. Oddsdóttir, H.Æ.; Hagrot, E.; Chotteau, V.; Forsgren, A. On dynamically generating relevant elementary flux modes in a metabolic network using optimization. *J. Math. Biol.* **2015**, *71*, 903–920. [[CrossRef](#)]
76. Leighty, R.W.; Antoniewicz, M.R. Dynamic metabolic flux analysis (DMFA): A framework for determining fluxes at metabolic non-steady state. *Metab. Eng.* **2011**, *13*, 745–755. [[CrossRef](#)]
77. Hebing, L.; Neymann, T.; Thüte, T.; Jockwer, A.; Engell, S. Efficient generation of models of fed-batch fermentations for process design and control. *IFAC-PapersOnLine* **2016**, *49*, 621–626. [[CrossRef](#)]
78. Hebing, L.; Neymann, T.; Engell, S. Application of dynamic metabolic flux analysis for process modeling: Robust flux estimation with regularization, confidence bounds, and selection of elementary modes. *Biotechnol. Bioeng.* **2020**, *117*, 2058–2073. [[CrossRef](#)]
79. Abbate, T.; de Sousa, S.F.; Dewasme, L.; Bastin, G.; Vande Wouwer, A. Inference of dynamic macroscopic models of cell metabolism based on elementary flux modes analysis. *Biochem. Eng. J.* **2019**, *151*, 107325. [[CrossRef](#)]
80. Maton, M.; Bogaerts, P.; Vande Wouwer, A. Selection of a Minimal Suboptimal Set of EFMs for Dynamic Metabolic Modelling. In Proceedings of the IFAC PapersOnLine 11th IFAC Symposium on Advanced Control of Chemical Processes, Venice, Italy, 13–16 June 2021.
81. Haag, J.; Vande Wouwer, A.; Remy, M. A general model of reaction kinetics in biological systems. *Bioprocess Biosyst. Eng.* **2005**, *27*, 303–309. [[CrossRef](#)]
82. Richelle, A.; Bogaerts, P. Systematic methodology for bioprocess model identification based on generalized kinetic functions. *Biochem. Eng. J.* **2015**, *100*, 41–49. [[CrossRef](#)]
83. Kaufman, D.E.; Smith, R.L. Direction choice for accelerated convergence in hit-and-run sampling. *Oper. Res.* **1998**, *46*, 84–95. [[CrossRef](#)]
84. De Martino, D.; Mori, M.; Parisi, V. Uniform sampling of steady states in metabolic networks: heterogeneous scales and rounding. *PLoS ONE* **2015**, *10*, e0122670. [[CrossRef](#)] [[PubMed](#)]
85. Bogaerts, P.; Rooman, M. DISCOPOLIS 2.0: A new recursive version of the algorithm for uniform sampling of metabolic flux distributions with linear programming. In Proceedings of the IFAC PapersOnLine 11th IFAC Symposium on Advanced Control of Chemical Processes, Venice, Italy, 13–16 June 2021.
86. Yasemi, M.; Jolicoeur, M. Modelling Cell Metabolism: A Review on Constraint-Based Steady-State and Kinetic Approaches. *Processes* **2021**, *9*, 322. [[CrossRef](#)]