

Algorithms to infer metabolic flux ratios from fluxomics data

Rafael Carreira and Miguel Rocha
CCTC, School of Engineering
Campus de Gualtar, Braga, Portugal
Email: {rafaelcc,mrocha}@di.uminho.pt

Silas G. Villas-Bôas
School of Biological Sciences
University of Auckland, New Zealand
Email: s.villas-boas@auckland.ac.nz

Isabel Rocha
CEB/ IBB, School of Engineering
Campus de Gualtar, Braga, Portugal
Email: irocha@deb.uminho.pt

Abstract—*In silico* cell simulation approaches based in the use of genome-scale metabolic models (GSMMs) and constraint-based methods such as Flux Balance Analysis are gaining importance, but methods to integrate these approaches with omics data are still greatly needed. In this work, the focus relies on fluxomics data that provide valuable information on the intracellular fluxes, although in many cases in an indirect, incomplete and noisy way. The proposed framework enables the integration of fluxomics data, in the form of ^{13}C labeling distribution for metabolite fragments, with GSMMs enriched with carbon atom transition maps. The algorithms implemented allow to infer labeling distributions for fragments/metabolites not measured and to build expressions for the relevant flux ratios that can be then used to enrich constraint-based methods for flux determination. This approach does not require any assumptions on the metabolic network and reaction reversibility, allowing to compute ratios originating from coupled joint points of the network. Also, when enough data do not exist, the system tries to infer ratio bounds from the measurements.

I. INTRODUCTION

Metabolic networks provide information about possible pathways for the conversion of nutrients into products, energy and reducing equivalents, but do not provide information about their actual fluxes [1]. Genome-scale metabolic models (GSMMs) are related to metabolic networks, but go one step further, supporting the simulation of the metabolic phenotype. Although the process is only semi-automatic, a number of GSMMs have been reconstructed for different organisms [2].

Metabolic Flux Analysis (MFA) uses GSMMs to calculate intracellular fluxes [3]. Indeed, taking into account reaction stoichiometry and reversibility provided by GSMMs, and imposing quasi-steady-state assumptions, linear constraints are defined, formulating a system to obtain the unknown fluxes. However, the majority of these systems are underdetermined and optimization methods need to be used, e.g. in Flux Balance Analysis (FBA) [4] a linear programming problem is solved to reach a flux distribution given an objective function. However, these approaches have some drawbacks, such as, configuring proper objective functions, alternative optimal solutions, reliance on cofactor balances and futile cycles [1].

Some of these problems can be attenuated when experimental measurements of metabolic fluxes are used, such as the rates of formation / consumption of some compounds, but these are rarely able to turn the system into a determined one.

^{13}C tracer experiments, where cells are fed with carbon labeled substrates, allow to perform more reliable analyses on intracellular fluxes [5]. A number of analytical tools can be used to perform the comprehensive determination of the ^{13}C labeling distribution of intracellular metabolites, including mass spectroscopy (MS) that will be the focus of this work. In metabolic networks, the carbon backbones of metabolites can be derived from alternative pathways. These experiments enable to track how the ^{13}C labeling patterns of the substrates are propagated throughout the network.

^{13}C -based approaches for MFA (^{13}C -MFA) can be classified as implementing global or local methods of inference. The former consists in nonlinear optimization problems that iteratively search for the optimum flux distribution that fits the observed data [6][7]. However, their complex formulations might hamper the interpretation of the results, and no direct evidence to a particular flux is given [8].

Metabolic flux ratio (METAFoR) analysis [9][10] is a local approach that uses probabilistic equations to interpret the ^{13}C -labeling pattern, by calculating the ratios of fluxes that produce a given metabolite through the different pathways. Since the calculation of these ratios is independent of the flux distribution [11], no knowledge is needed either on the rates of uptake and formation of external metabolites, or on the biomass composition [12]. From the metabolic flux ratios, constraints can be configured to reduce the degrees of freedom of the systems in stationary MFA.

Usually, flux ratios are calculated from manually derived equations [13][11][14]. In previous work, a framework has been developed [12] that aims to infer an equation system to constrain the flux distribution to a low-dimensional convex set. However, this framework still shows some limitations since it depends on simplifying assumptions about the network topology, as it resorts only to linear optimization techniques, or flux directionality, and it does not provide a general-purpose way of creating bounds over flux ratios.

In this work, a novel automated approach is proposed, that is independent of any assumptions on the metabolic network and on the reactions reversibility. This framework generates all possible equations to calculate metabolic flux ratios for a given GSMM, taking as input a set of intact carbon fragments of internal metabolites for which relative abundances have been measured (or determined from the abundance of the amino

acids for which they are precursors [10]). Moreover, when not enough labeling measurements are given to the system, equations are built to estimate bounds for the ratios. In the cases where the generated equations are exact and linear, the ratios can be algebraically calculated. Otherwise, a Differential Evolution (DE) method is configured to estimate these ratios. The obtained ratios can be useful to set up constraints for existing constraint-based flux analysis methods [15], [16].

II. DEFINITIONS AND PROBLEM STATEMENT

A. Mass isotopomer distribution vectors

When ^{13}C -labeled substrates are fed into cells, the labeling of the substrate is propagated throughout the network, and the compounds show a labeling pattern that is the consequence of how the carbons are distributed. A molecule with n carbon atoms can have 2^n different labeling patterns, due to the possible configurations in the number and position of labeled atoms. Each labeling pattern defines a positional isotopomer, while the set of isotopomers with the same number of labeled atoms defines the mass isotopomers of the molecule.

Based on these mass isotopomers, for an n carbon-sized molecule, it is possible to build a probability vector, named mass isotopomer distribution vector (MDV) with $n+1$ positions, where the i -th position of the vector corresponds to the probability of the molecule being labeled on i carbon atoms. Using experimental data from MS, the probabilities are assigned from the fractional abundance of the m_{+i} mass of the fragment, where the m_{+0} corresponds to the mass of the lightest isotopomer, and the remaining are the abundance of molecules with higher masses [11].

The experimental data are often obtained from techniques that detect fragments instead of entire molecules, such as, gas chromatography coupled with mass spectrometry (GC-MS), where a set of ion fragments from different parts of the carbon skeleton of the molecules are detected. It is, therefore, important to introduce the notation of a fragment as $A_{i,j}$, representing a fragment of compound A that starts in its i -th and ends in its j -th carbon atoms. When the measured data correspond to aminoacids, after some operations to correct for the natural isotope abundances [17], it is possible to obtain the same type of vectors for its precursors [11].

In the following, \mathcal{F} will denote the set of fragments that have been measured, i.e., that contain an associated MDV. Note that the proposed framework is designed to work over this set, while other fragments and metabolites are not measured and their MDVs can only be inferred (if enough data are available).

B. Carbon atoms mapping

To perform the tasks described next, models with additional information on how the carbon atoms are transferred in each reaction are necessary [18]. These models contain the mapping of each reaction, which consists in a bijection of the reactants' to the products' carbon atoms [19], elucidating how the labeling patterns in the carbon skeleton are propagated through all the metabolites in the network.

The model enriched with carbon atom mapping can be represented as a graph $\mathcal{N} = (\mathcal{M}, \mathcal{T}, \mathcal{I})$, where \mathcal{M} is the set of nodes associated to the metabolites, \mathcal{T} is the set of edges representing carbon transitions between the reactants and products of the corresponding reactions and \mathcal{I} indicates the set of transitions that are irreversible.

C. Inferring mass isotopomer distribution vectors

This work aims at determining ratios that reflect the contribution of different paths to the labeling distribution of metabolite fragments in the network. Given each fragment in \mathcal{F} , it is possible to navigate over \mathcal{N} , to check if this fragment is a product reached from distinct paths in the graph and, thus, allows to calculate a flux ratio for this metabolite.

Here, a path p to a product fragment is defined as $\rho_p = (\mathcal{T}_p, \mathcal{F}_p)$ with $\mathcal{T}_p \subseteq \mathcal{T}$ and $\mathcal{F}_p \subseteq \mathcal{F}$, which, through a series of transitions \mathcal{T}_p associated to a set of intermediate metabolites in \mathcal{M} , connects a set of measured fragments \mathcal{F}_p to the product (also in \mathcal{F}). Let us consider \mathcal{P} as the set of paths originating a metabolite fragment Z . If $|\mathcal{P}| \geq 2$, an expression to infer the MDV of Z can be built using the algorithms that will be explained in sections III and IV. Since there are at least two alternative paths to Z , the expression will contain ratio variables that define the fraction of the labeling contribution of each path. The following step is to reach the values for these variables, a process that will be explained in section V.

III. LINKING SOURCE FRAGMENTS TO PRODUCTS

When searching for the set of fragments in \mathcal{F} that originate a target fragment Z , different scenarios may occur. If \mathcal{P} is the set of paths originating the carbon backbone of Z , the simplest case occurs when $|\mathcal{P}| = 1$ (ρ_1) and $|\mathcal{F}_1| = 1$ (Fig. 1 A). Therefore, the carbons that contribute to Z can be traced by following the carbon transfers on the transitions in ρ_1 and the MDV of Z is equal to the MDV of the source fragment.

However, if $|\mathcal{F}_1| > 1$ (i.e. at least one transition in \mathcal{T}_1 has more than one substrate), the combination of the labeling of all the fragments of \mathcal{F}_1 must be considered (Fig. 1 B).

Please note that, in the following expressions, the MDV of a metabolite fragment Z is simply given by the representation of the fragment and $Z(i)$ is used to denote the value of the MDV in the position i , i.e. the probability of Z being labeled at i carbon atoms. Thus, the MDV of Z with n carbons, that is a combination of all the fragments in $\mathcal{F}_1 = \{S^1, S^2, \dots, S^m\}$, is a column vector, where each position k is obtained as:

$$Z(k) = \sum_{\substack{\text{all combinations} \\ \text{where } \sum_{i=1}^m l_i = k}} \left(\prod_{i=1}^m S^i(l_i) \right), k = 0, \dots, n \quad (1)$$

where l_i is the labeling of S^i in the combination.

In the case where $|\mathcal{P}| > 1$ (Fig. 1 C), the labeling pattern of Z can be influenced by a fraction τ_i of each converging path ρ_i . These fractions define ratios that are related to the activity of the different pathways regarding their flux contribution to

the product. In such scenario, the proportions of fluxes in each path are used to calculate the MDV of the product, together with the MDVs of the sources \mathcal{F}_i for each ρ_i :

$$Z = \sum_{i=1}^{|\mathcal{P}|} \tau_i \cdot MDV_{\rho_i}, \sum_{i=1}^{|\mathcal{P}|} \tau_i = 1 \quad (2)$$

where τ_i is the ratio of Z that is derived from the pathway ρ_i that has a measured (or calculated) MDV given by MDV_{ρ_i} .

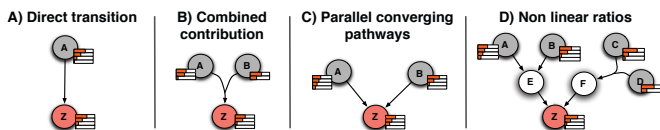


Fig. 1. Labeling contribution (orange bars) of the precursors (grey circles) to their products (magenta circles). The arrows represent the carbon atom transitions between the reagents and the products of each reaction.

However, the scenario becomes more complex when, in the search for a branch with a converging ratio, it is not possible to obtain the MDV of a fragment source of the pathway in \mathcal{F} , without considering another ratio (Fig. 1 D). In such cases, the unknown ratios will be multiplied by each other and a nonlinear expression is obtained for the MDV of Z :

$$Z = \tau_1 \cdot \tau_{1.1} \cdot A + \tau_1 \cdot \tau_{1.2} \cdot B + \tau_2 \cdot (C \times D) \quad (3)$$

where τ_1 , τ_2 , $\tau_{1.1}$ and $\tau_{1.2}$ are the ratios of the converging pathways to Z , $\tau_1 + \tau_2 = 1$, $\tau_{1.1} + \tau_{1.2} = 1$ and \times represents the operator defined by expression 1 above.

IV. CALCULATING RATIO EXPRESSIONS

The above expressions can be combined to translate how the MDV of a product is influenced by its precursors. Each fragment in \mathcal{F} is selected as a candidate product for a ratio expression and the process searches on \mathcal{N} for the fragments in \mathcal{F} that can contribute to its labeling pattern.

A. Recurrent sub-expressions

During the procedure, the sources are not always found in the nodes that are immediately precedent to the product. Indeed, almost all expressions are built upon pathways where $|\mathcal{T}_p| \gg 1$. In each iteration, every reactant fragment in each transition that originated the current node, i.e. an intermediate fragment, will be recursively selected as a product node in the following iterations. This process can be seen as a search for the sub-expressions of the intermediates that form the global expression to the original product fragment.

Metabolic networks are complex and highly connected, so the carbon atoms transitions between given metabolites are also complex. The same carbon skeleton can contribute for different metabolites of the network, and it can even contribute to the same fragment from alternative pathways, as illustrated in Fig. 2 A. Therefore, when searching for the expression that enables to obtain a fragment from its sources, it is possible to find metabolites for which the expressions for its fragments had been previously derived. In such cases, if the expression has already been found, it is not necessary to repeat the procedure of calculating its MDV. Thus, a register that keeps the history of the searches is kept and updated whenever a new

fragment is found and its expression is built. If a fragment has been already visited in some previous iteration, or in the search for a different product of \mathcal{F} , the expression of the fragment is retrieved from the register and attached to the expression of the selected product. Hence, the expression to determine the MDV of $Z_{2,5}$ can be written as:

$$\begin{cases} Z_{2,5} = \tau_1 \cdot E_{2,5} + \tau_2 \cdot (\tau_{2.1} \cdot D_{1,4} + \tau_{2.2} \cdot G_{2,5}) \\ E_{2,5} = \tau_{1.1} \cdot A_{2,5} + \tau_{1.2} \cdot (B_{2,3} \times C_{1,2}) \end{cases} \quad (4)$$

B. Undetermined ratios

The expressions obtained are totally dependent on \mathcal{F} . Therefore, the sources of a given branch might not be found in the network. When searching for a branch, if there are no measurements that enable to calculate the vectors that originate all the metabolites that are reached, it will not be possible to find all the sources of the labeling contributions from that branch. To guarantee that the search process terminates, a stopping criteria is defined allowing a maximum number of nodes to be visited. When the limit is reached, it is assumed that it is not possible to find sources belonging to \mathcal{F} on the path and the corresponding part of the expression cannot be used to calculate the MDV of the product. However, it should not be assumed that there is no influence of that path on the labeling pattern of the product. Hence, one can either dismiss the expression, if the degree of uncertainty is too high (e.g. too many unknown sources) or use the expression to estimate partial bounds for the ratios using the pathways for which it is possible to find sources belonging to \mathcal{F} .

In Fig. 2 B the search is driven to a path where no fragment of \mathcal{F} is reached. Therefore, the contribution of the path is compromised by this missing vector and the MDV of Z can not be accurately calculated. If the branch originating from H is not accounted in the expression, it can be written as:

$$Z_{2,5} = \tau_1^* \cdot (\tau_{1.1} \cdot A_{2,5} + \tau_{1.2} \cdot B_{1,4}) + \tau_2^* \cdot (\tau_{2.1}^* \cdot C_{2,5} + \tau_{2.2}^* \cdot D_{2,5}) \quad (5)$$

where τ_1^* , $\tau_{2.1}^*$ and $\tau_{2.2}^*$ are the upper bounds of the ratios to $Z_{2,5}$ originating from v_1 , v_5 and v_6 , respectively, and τ_2^* is the lower bound of the ratio from v_4 .

C. Carbon mapping cycles

Occasionally, a fragment is reached again on a further iteration, before any fragment in \mathcal{F} is found. In such cycles, it will not be possible to determine the corresponding part of the expression due to missing fragment sources. Note that a cycle does not refer to repeated metabolites, but to fragments.

In Fig. 2 C, the process is driven to visit the fragment $G_{2,3}$ when searching for the sources of the fragment $F_{2,5}$, but it is visited again, before finding the sources to the fragment. Therefore, an expression to calculate the exact flux ratios cannot be found, but some ratio bounds can be estimated:

$$Z_{2,5} = \tau_1^* \cdot A_{2,5} + \tau_2^* \cdot (\tau_{2.1}^* \cdot C_{2,5} + \tau_{2.2}^* \cdot (D_{1,2} \times B_{2,3})) \quad (6)$$

where τ_1^* and $\tau_{2.1}^*$ are the upper bounds of the ratios of reactions v_1 and v_3 , respectively, and $\tau_{2.2}^*$ and $\tau_{2.2}^*$ the lower bounds of the ratios for reactions v_2 and v_5 , respectively.

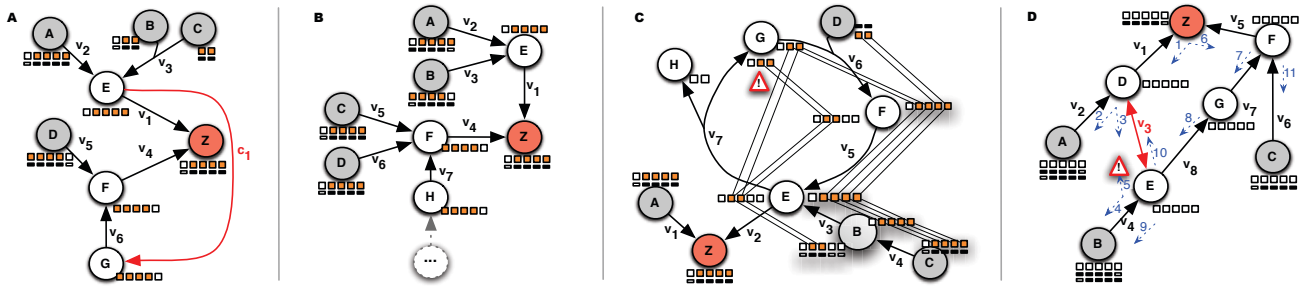


Fig. 2. Small network excerpts showing different cases when tracing the labeling contribution to a product (Z) fragment. The orange squares stand for the contiguous fragments that are visited during the search process, and the sequence of white / black rectangles identifies the measured fragments. In A, the red arrow indicates a reaction that originates a repeated search for the same fragment of G . In B, “...” illustrates a branch in which it is not possible to find measured sources that originate $Z_{2,5}$. In C, the fragment $G_{2,5}$ is visited twice and the thin lines connect the carbon atoms that are being transferred between the fragments. In D, blue arrows indicate the direction and order taken in the process, with no previous assumption on the direction of the reversible reaction (the red arrow). The coloured carbons in D were omitted, since the carbon fragments of some metabolites vary during the iterations of the process.

D. Reversible reactions

When the model contains information on the reversibility of the reactions, each reaction $v_i \notin \mathcal{I}$ is decomposed on two irreversible reactions in both directions. However, a reaction cannot contribute simultaneously, on both directions, to the labeling of the metabolites involved in its pathways. When there are no insights on which direction a certain reaction is taking place in a physiological context, if every possibility is taken into account, the problem would be combinatorial and many different solutions would be returned.

In Fig. 2 D, $v_3 \notin \mathcal{I}$ and the carbon transitions can occur on either direction. If the procedure progresses in the order of the blue arrows, initially v_3 will be visited to search for the sources that originate $Z_{1,4}$. Afterwards, the process will proceed to find the sources of $E_{1,4}$ and, in the reverse way, v_3 will be selected. However, since v_3 has already been selected on the direction E to D , it does not make sense to select it in the reverse direction. Thus, for the same branch, if a reaction has been previously selected on one direction, the reverse is not considered. Nevertheless, when the search is performed to find the sources to $Z_{1,4}$, across the path originating from v_5 , the reaction v_3 will be reached again, but on a different direction of the one kept. If there is no insight on the proper direction of the reaction, both directions are considered. So, the direction of the reactions is locally defined only for a given branch during the process, but both directions can be used in the global calculation of the expressions.

Here, instead of generating all the possibilities for the expressions, no filter or stopping criterion is defined. Indeed, the two directions are interpreted as two different reactions, and a generic expression is built. To solve it, the proper direction of the reactions must be selected, either from the experimental measurements or from knowledge on the physiological conditions of the problem. Subsequently, the excerpts of the paths that are derived from the invalid directions, are removed by setting the corresponding ratios to zero.

V. DETERMINATION OF METABOLIC FLUX RATIOS

After obtaining the expressions that enable to calculate the MDVs for the fragments in \mathcal{F} , a set of unknowns is defined for the flux ratios. To compute the values of these ratios, two scenarios can occur. The simplest case is when the expression

to calculate a product fragment is linear. In the other case, the expressions are not linear due to the occurrence of products common in different flux ratios.

The former relates to expressions that encompass only one level of ratios, with no multiplication of ratio variables. In such cases, it is possible to calculate the flux ratios by rearranging the equation 2. Since $\tau_m = 1 - \tau_1 - \tau_2 - \dots - \tau_{m-1}$, the following system for the first $m - 1$ flux ratios can be written as:

$$\begin{bmatrix} MDV_1 - MDV_m \\ MDV_2 - MDV_m \\ \dots \\ MDV_{m-1} - MDV_m \end{bmatrix} \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_{m-1} \end{bmatrix} = [MDV_Z - MDV_m] \quad (7)$$

where $m = |\mathcal{P}|$, the left-hand side is an $n \times m - 1$ matrix A covering the subtraction of the MDVs of the first $m - 1$ sources by the source m , τ is a vector of the first $m - 1$ ratios and the right-hand side $n \times 1$ vector (b) corresponds to the subtraction of the MDV of the product and the source n .

The other case occurs when pathways lead to an intermediate node that does not belong to \mathcal{F} . Here, the problem tries to find the vector τ that minimizes the difference between the calculated and measured MDVs of the product without using any gradient method. Instead, a Differential Evolution (DE) method is configured to solve the optimization problem formulated below, by iteratively trying to improve a candidate solution with regard to the objective function and constraints, defined as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{|\mathcal{D}|} \sum_{j=1}^{s_i} \left(\frac{MDV_{c,i}(j) - MDV_{m,i}(j)}{MDV_{m,i}(j)} \right)^2 \\ \text{subject to:} \quad & \sum_{j=1}^{s_i} MDV_{c,i}(j) = 1, \quad \forall MDV_i \in \mathcal{D} \\ & \sum_{x=1}^{s_l} \tau_{l,x} = 1, \quad \forall l \in \mathcal{R} \end{aligned}$$

where $MDV_{c,i}(j)$ and $MDV_{m,i}(j)$ are the calculated and measured fractional abundances of the $m+j$ mass of the fragment i , respectively, $\mathcal{D} \subset \mathcal{F}$ the products for which the DE is being solved, s_i is the length of MDV_i and s_l is the number of elements in the subset l in $\mathcal{R} = \{\{\tau_{1,1}, \dots, \tau_{1,s_1}\}, \dots, \{\tau_{l,1}, \dots, \tau_{l,s_l}\}\}$, that encompass all the subsets of ratios of pathways originating the same intermediate fragment. The values of $m_{c,i+j}$ are calculated from the obtained expressions to the corresponding products, after the assignment of the flux ratios.

Fig. 3 illustrates how a candidate solution is encoded and used to solve the problem. A solution is a vector that encodes real values corresponding to the ratios in the expressions. Although the expressions refer to different product fragments, since some flux ratios are shared, all ratios are encoded in the same solution. In fact, when ratios are shared, all the expressions that use these ratios will contribute to their computation, making the estimation more accurate. However, the expressions that do not have intersections between its set of ratios are formulated and solved independently, meaning that each solution used to the calculation of the ratios is local to the corresponding expression, and the ratios are not globally calculated in a single solution.

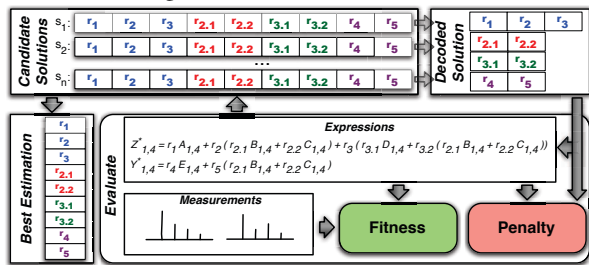


Fig. 3. Differential Evolution, including solution encoding and evaluation.

In each iteration, a candidate solution is evaluated by assigning the values encoded in the solution to the corresponding ratios in the expressions. The expressions are solved and the calculated MDVs are compared with the measured ones. Since the objective is to minimize this difference, a fitness value to each candidate solution will be proportional to the squared errors in the estimation. To make the method respect the constraints, a penalty function is configured to penalize solutions that originate a calculated MDV for which the sum of its elements is not equal to one, and solutions that contain groups of ratios (in the same level, corresponding to different paths leading to the same intermediate node) for which the sum is different from one. Therefore, a metric is defined to weight the evaluation of the candidate solutions with respect to its fitness and penalty values.

VI. CONCLUSION

The framework presented here allows the flexible and automatic computation of metabolic flux ratios for performing tasks of ^{13}C metabolic flux analysis. Based on a set of input fragments, that have an associated MDV obtained from ^{13}C labeling experiments, and on a graph that represents the topology of the metabolic network and how the carbon atoms are transferred between reactants and products, this approach traces the carbon atoms through the different pathways of the network. As its main results it is able to compute MDV expressions for fragments not measured and to build expressions that enable to obtain the relevant flux ratios.

This approach does not require any assumptions on the metabolic network and reversibility of the reactions, and allows to compute ratios originating from coupled joint points of the network. Since the measured data influence the shape of the derived equations, when enough data do not exist, the

system tries to infer ratio bounds from the measurements. The computed ratios can be used to generate metabolic flux ratio constraints to be used within existing approaches to perform constraint-based flux analysis methods.

ACKNOWLEDGMENT

The work is partially funded by ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT (Portuguese Foundation for Science and Technology) within project ref. COMPETE FCOMP-01-0124-FEDER-015079. RC's work is funded by a PhD grant from the Portuguese FCT (ref. SFRH/BD/66201/2009).

REFERENCES

- [1] G. Stephanopoulos *et al.*, *Metabolic engineering: principles and methodologies*. Academic Press, 1998.
- [2] C. S. Henry, *et al.*, "High-throughput generation, optimization and analysis of genome-scale metabolic models," *Nature biotechnology*, vol. 28, no. 9, pp. 977–982, 2010.
- [3] J. J. Vallino *et al.*, "Metabolic flux distributions in corynebacterium glutamicum during growth and lysine overproduction," *Biotechnology and Bioengineering*, vol. 41, no. 6, pp. 633–646, 1993.
- [4] J. Edwards *et al.*, "In silico predictions of escherichia coli metabolic capabilities are consistent with experimental data," *Nature biotechnology*, vol. 19, no. 2, pp. 125–130, 2001.
- [5] C. R. Malloy *et al.*, "Evaluation of carbon flux and substrate selection through alternate pathways involving the citric acid cycle of the heart by ^{13}C NMR spectroscopy," *Journal of Biological Chemistry*, vol. 263, no. 15, pp. 6964–6971, 1988.
- [6] W. Wiechert *et al.*, "A universal framework for ^{13}C metabolic flux analysis," *Metabolic engineering*, vol. 3, no. 3, pp. 265–83, Jul. 2001.
- [7] J. D. Young *et al.*, "An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis," *Biotechnology and bioengineering*, vol. 99, no. 3, pp. 686–699, 2008.
- [8] M. Emmerling *et al.*, "Metabolic Flux Responses to Pyruvate Kinase Knockout in Escherichia coli," *Society*, vol. 184, pp. 152–164, 2002.
- [9] U. Sauer *et al.*, "Metabolic flux ratio analysis of genetic and environmental modulations of Escherichia coli central carbon metabolism," *Journal of bacteriology*, vol. 181, no. 21, pp. 6679–88, Nov. 1999.
- [10] T. Szyperski, "Biosynthetically Directed Fractional ^{13}C -labeling of Proteinogenic Amino Acids," *European Journal of Biochemistry*, vol. 232, no. 2, pp. 433–448, 1995.
- [11] E. Fischer *et al.*, "Metabolic flux profiling of Escherichia coli mutants in central carbon metabolism using GC-MS," *European Journal of Biochemistry*, vol. 270, no. 5, pp. 880–891, 2003.
- [12] A. Rantanen *et al.*, "An analytic and systematic framework for estimating metabolic flux ratios from ^{13}C tracer experiments," *BMC bioinformatics*, vol. 9, p. 266, Jan. 2008.
- [13] U. Sauer *et al.*, "Metabolic fluxes in riboflavin-producing *Bacillus subtilis*," *Nature Publishing Group*, vol. 15, pp. 448–452, 1997.
- [14] L. M. Blank *et al.*, "Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts," *FEMS yeast research*, vol. 5, no. 6-7, pp. 545–58, Apr. 2005.
- [15] N. Zamboni *et al.*, "FiatFlux—a software for metabolic flux analysis from ^{13}C -glucose experiments," *BMC bioinformatics*, vol. 6, p. 209, 2005.
- [16] M. J. McNulty *et al.*, "Genome-scale modeling using flux ratio constraints to enable metabolic engineering of clostridial metabolism in silico," *BMC systems biology*, vol. 6, no. 1, p. 42, 2012.
- [17] W. A. van Winden, *et al.*, "Correcting mass isotopomer distributions for naturally occurring isotopes," *Biotechnology and bioengineering*, vol. 80, no. 4, pp. 477–9, Dec. 2002.
- [18] P. Ravikirithi, *et al.*, "Construction of an E. Coli genome-scale atom mapping model for MFA calculations," *Biotechnology and bioengineering*, vol. 108, no. 6, pp. 1372–1382, 2011.
- [19] M. Latendresse *et al.*, "Accurate Atom-Mapping Computation for Biochemical Reactions," *Journal of chemical information and modeling*, vol. 52, no. 11, pp. 2970–2982, 2012.