



HAL
open science

Stochastic simulation algorithm for isotope-based dynamic flux analysis

Quentin Thommen, Julien Hurbain, Benjamin Pfeuty

► **To cite this version:**

Quentin Thommen, Julien Hurbain, Benjamin Pfeuty. Stochastic simulation algorithm for isotope-based dynamic flux analysis. *Metabolic Engineering*, 2022, 75, pp.100 - 109. 10.1016/j.ymben.2022.11.001 . hal-04012821

HAL Id: hal-04012821

<https://hal.science/hal-04012821v1>

Submitted on 12 Jan 2024

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

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

1 **Stochastic simulation algorithm for isotope-based dynamic flux analysis**

2 Quentin Thommen,¹ Julien Hurbain,² and Benjamin Pfeuty²

3 ¹*Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, UMR9020-UI277 -*
4 *CANTHER - Cancer Heterogeneity Plasticity and Resistance to Therapies, F-59000 Lille,*
5 *France*

6 ²*Univ. Lille, CNRS, UMR 8523 - PhLAM - Physique des Lasers Atomes et Molécules,*
7 *F-59000 Lille, France*

8 (*Electronic mail: quentin.thommen@univ-lille.fr)

9 (Dated: 4 October 2022)

10 Carbon isotope labeling method is a standard metabolic engineering tool for flux quantifi-
11 cation in living cells. To cope with the high dimensionality of isotope labeling systems,
12 diverse algorithms have been developed to reduce the number of variables or operations in
13 metabolic flux analysis (MFA), but lacks generalizability to non-stationary metabolic con-
14 ditions. In this study, we present a stochastic simulation algorithm (SSA) derived from the
15 chemical master equation of the isotope labeling system. This algorithm allows to com-
16 pute the time evolution of isotopomer concentrations in non-stationary conditions, with the
17 valuable property that computational time does not scale with the number of isotopomers.
18 The efficiency and limitations of the algorithm is benchmarked for the forward and in-
19 verse problems of ¹³C-DMFA in the pentose phosphate pathways. Overall, SSA consti-
20 tute an alternative class to deterministic approaches for metabolic flux analysis that is well
21 adapted to comprehensive dataset including parallel labeling experiments, and whose lim-
22 itations associated to the sampling size can be overcome by using Monte Carlo sampling
23 approaches.

24 **Keywords:** Metabolic flux analysis, Flux balance analysis, Metabolism, Metabolic net-
25 work model, Stable-isotope tracers, Systems biology

1. INTRODUCTION

Isotope tracing experiments have been developed to quantify fluxes in biochemical networks (Stephanopoulos, 1999). A typical carbon-13 labeling experiment metabolizes a labeled substrate, such as [1- ^{13}C]glucose, tracks the propagation of the label on metabolites by nuclear magnetic resonance (NMR) or mass spectrometry (MS) methods and estimates metabolic fluxes by various methods including ^{13}C -MFA (Niederführ, Wiechert, and Nöh, 2015; Allen and Young, 2020; Antoniewicz, 2021). Despite its limitations, ^{13}C -MFA remains the gold standard method in metabolic engineering for accurate and precise quantification of fluxes in living cells (Crown and Antoniewicz, 2013). Currently, the most efficient algorithms are all based on an advanced decomposition method using elementary metabolic units (EMUs) developed in 2007 by Antoniewicz et al (Antoniewicz, Kelleher, and Stephanopoulos, 2007). Nevertheless, one of the limitations of the classical metabolic flux analysis (MFA) method is the requirement of a metabolic isotopic steady state. Flux analysis methods that focus on estimating non-stationary metabolic fluxes are referred to as dynamic MFA (DMFA) (Leighty and Antoniewicz, 2011), or ^{13}C dynamic MFA methods (^{13}C -DMFA) methods (Antoniewicz, 2015a). Despite pioneering works (Antoniewicz *et al.*, 2007; Wahl, Nöh, and Wiechert, 2008) initiated more than one decade ago, little progress has been made since (Antoniewicz, 2021). Current computational methods use a deterministic modeling framework by solving EMU balance rate equations where dynamic flux parameters are modeled with B-splines (Quek *et al.*, 2020; Ohno *et al.*, 2020). Computational tractability of such method depends on the EMU system size that can be very large due to the interplay of elaborated labeling protocols (Lewis *et al.*, 2014; Antoniewicz, 2015b; Jacobson *et al.*, 2019; Dong *et al.*, 2019; Allen and Young, 2020) and complex biotic reactions (Selivanov *et al.*, 2004).

In this paper, we present a different class of method that simulates isotope propagation in non-stationary metabolic systems by a Stochastic Simulation Algorithm (SSA). We test the method in the metabolic subsystem comprising glycolytic and PPP pathways where complex carbon rearrangements occur due to biotic reactions in the nonoxidative PPP and where ^{13}C labeling have been extensively applied to infer metabolic flux (Kuehne *et al.*, 2015; Bouzier-Sore and Bolaños, 2015; Creek *et al.*, 2015; Diaz-Moralli *et al.*, 2016; Lee *et al.*, 2019). The main idea is to represent the population of isotopomers of a chemical species by a sample of finite size, proportional to the species concentration, and to use the standard rules of stochastic chemical kinetics to propagate the marker. When a reaction occurs, the isotopomers associated to the reactants are

57 randomly selected in the corresponding samples, the rearrangement is performed, and the products
58 are added to the corresponding samples (Figure 1) The algorithm somehow mimics the discrete
59 and stochastic processes of enzymatic reactions as it occurs in cells, but remains restricted to a
60 small sample of metabolite species for the sake of computational efficiency. At each time step, the
61 samples represent the population of the isotopomers of each variable from which one can compute
62 mass isotopomer distribution for comparing with experimental data. The proposed algorithm is
63 simple to implement, fast, visual, and above all its computation time depends very little on the
64 chain length, which makes it an algorithm also adapted to parallel labeling (^{13}C , ^2H , ^{15}N , ^{18}O ,
65 etc).

66 2. RESULTS

67 2.1. Stochastic Simulation Algorithm (SSA)

68 The propagation of labeled atoms through a biochemical network is here described by a sam-
69 pling approach. The representation of the isotopomer distribution of each chemical species in the
70 network is computed using a finite sample size proportional to its concentration. A user defined
71 parameter Ω corresponds to a reference concentration. For example, a value of $\Omega = 1000 \text{ c}/\mu\text{M}$
72 indicates that a concentration of $1 \mu\text{M}$ is represented by 1000 copies of the chemical species, each
73 copy corresponding to a different isotopomer.

74 The fluxes of chemical reactions are determined by mathematical functions that can be ei-
75 ther linked to the species concentrations in the framework of chemical kinetics, or described by
76 phenomenological functions depending on time, or by constant functions in the case of station-
77 ary flux condition. The flux value determines the time interval between two occurrences of the
78 corresponding chemical reactions. When one occurs, the reactants are taken randomly from the
79 corresponding samples, the rearrangement of the atoms is done according to the reaction's rule,
80 and the products are added to the corresponding samples. In this way, the labeling propagates
81 through the chemical reaction network; at a given date, the sample of each species is populated
82 with different isotopomers and represents the isotopomer distribution.

83 Such rules are formalized within the framework of the chemical master equation once two
84 new tools are defined, the isotopomer index and addressing operators (sec. 3.1). Chemical master
85 equation describes the temporal evolution of the isotopomer fraction. From the chemical master

Stochastic method for isotope labeling systems

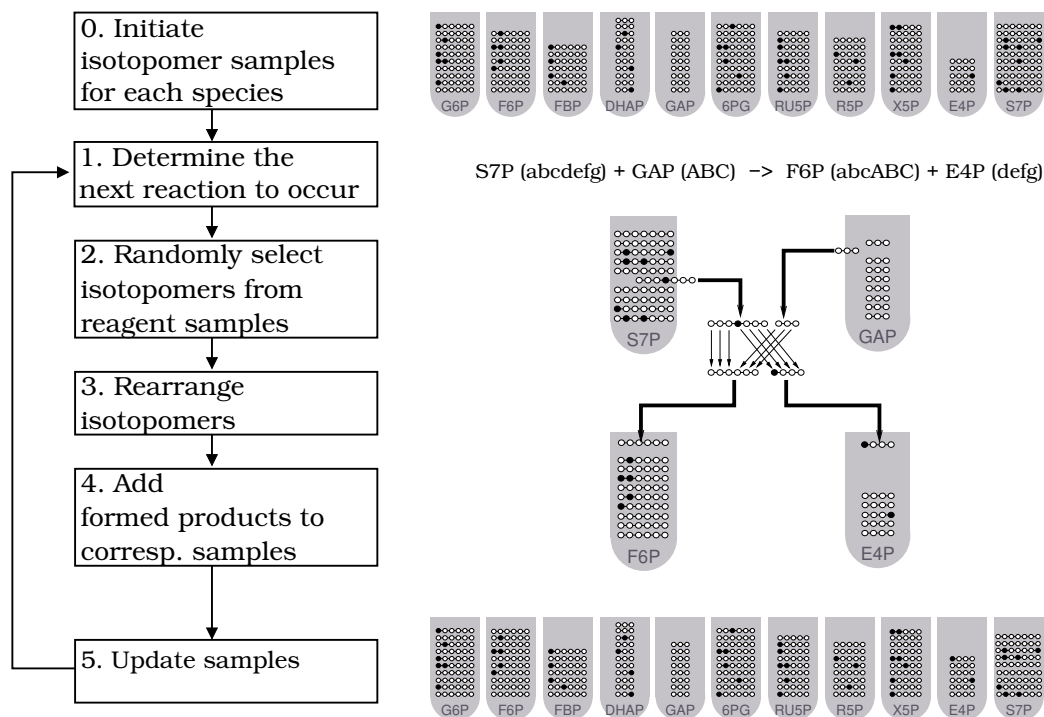


FIG. 1. General principle of the stochastic simulation algorithm (left part) and example of one iteration realization (right part). For each chemical species, the diversity of isotopomers is represented by a sample. The time evolution consists in determining the next reaction from the flux values, and then performing the reaction on reagents randomly selected in the corresponding samples. The products formed by internal rearrangement are added to the corresponding samples. The sample population is thus updated before the next reaction. The right side of the figure shows one iteration of the SSA where the next reaction to occur is the one mediated by the transaldolase.

86 equation, one can derive both a deterministic simulation algorithm (DSA) (see Sec. 3.2) and a
 87 stochastic simulation algorithm (SSA) (see Sec. 3.3). The DSA is not an efficient algorithm since
 88 it has as many variables as possible isotopomer, it is a "brute force" algorithm serving here as a
 89 control for the SSA outputs.

90 An example of stochastic simulation is given in Figure 2 and for the upper glycolytic pathway
 91 combined with the pentose phosphate pathway. To determine the fluxes, the mass action law is
 92 here used with unitary kinetic parameters (Table 1). At the initial time, the metabolic system is
 93 fed with labeled glucose (50% of [1-¹³C]glucose and 50% of [2-¹³C]glucose in (Kuehne *et al.*,
 94 2015)), and at the same time, is perturbed by a two-fold increase of the glucose intake rate. If
 95 $\Omega = 100 \text{ c}/\mu\text{M}$ is used for SSA, the Figure 2 (and the corresponding video) only represents one

Stochastic method for isotope labeling systems

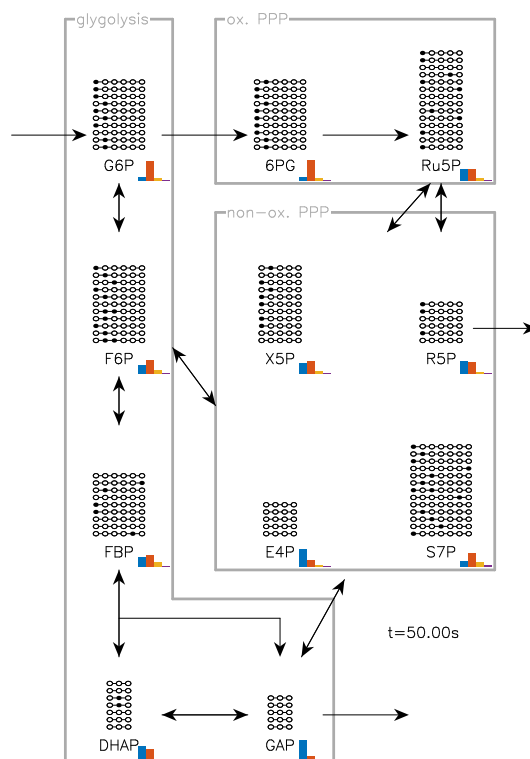


FIG. 2. Atom rearrangements in metabolic reactions. The upper glycolytic pathway supplemented by the pentose phosphate pathway provides an example of isotope labeling network. For each chemical species (G6P, F6P, FBP, DHAP, GAP, 6PG, Ru5P, X5P, R5P, E4P, S7P) subsamples of the isotopomer samples are displayed as a chain of unlabeled (open circle) and labeled carbons (close circle). The number of displayed carbons chains is proportional to the species concentration. Mass isotopomer histogram is also displayed ($m+0$ in black, $m+1$ in red, $m+2$ in green, and $m+3$ in blue). The Figure illustrates the configuration 50 s after the labeling introduction that also corresponds to the perturbation of the metabolic system. The associated video provides a full dynamical picture.

96 element out of 20 from each sample, for the sake of visualization.

97 The Figure 3 represents the evolution of the concentration and mass isotopomer obtained with
 98 the SSA (point) and the DSA (continuous line), thus depicting the accurate trends of isotopomer
 99 trajectories generated with SSA. The stochastic fluctuations of the mass isotopomers induced by
 100 the SSA are only due to the random selection of the reagents in the sample. The variance of these
 101 fluctuations is thus equal to the Ω profile and the mass isotopomer concentration. The determina-
 102 tion is thus all the more precise as the mass isotopomers are abundant. It is thus possible to reduce
 103 these fluctuations in two different ways, either by increasing the value of Ω , or by proceeding to a

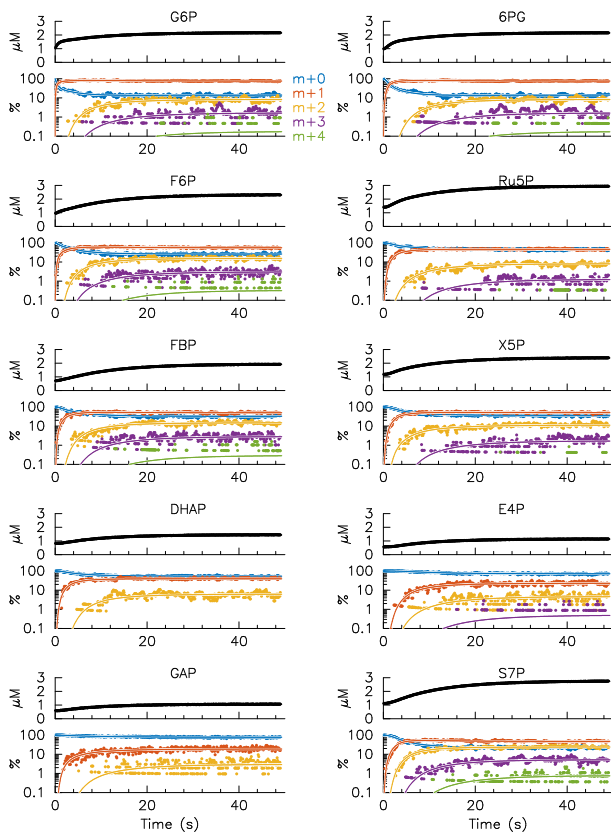


FIG. 3. Concentration and mass isotopomer dynamics in nonstationary conditions. Stochastic trajectories computed with SSA (dots) are compared with deterministic trajectories computed with DSA for control (solid lines), and corresponds to calculation presented in Figure 1 (same network, same condition). For each chemical species, the upper plot displays the concentration in μM whereas the bottom plot displays the mass isotopomers in percent ($m+0$ in black, $m+1$ in red, $m+2$ in green, and $m+3$ in blue, $m+4$ violet).

104 temporal smoothing of the stochastic evolution.

105 2.2. SSA Computational Performance

106 SSA computation time depends on both the number of chemical reactions and the execution
 107 time of each reaction. As an example, the computation cost necessary to simulate data of the
 108 Figure 2 corresponds to 224 463 reactions carried out in 62 ms for the SSA, and 2892 right-hand-
 109 side evaluation in 2100 ms for the DSA (advanced Runge-Kutta-Fehlberg method is used) using a
 110 Intel(R) Core(TM) i5-6300U CPU at 2.40GHz without parallelization. In the SSA, the number of
 111 chemical reaction occurrences can be approximated by the product $T \nu \Omega N$ where N is the number

112 of chemical reactions in the network, v the typical flux values and T the time interval. The number
113 of reactions does not depend on the number of isotopomer per species or, equivalently, on the chain
114 length representing the chemical species. The time to perform a reaction depends only slightly on
115 the chain length l thus almost not depend on the number of isotopomer. In our implementation,
116 the computation time for one reaction varies as $1 + l/15$; so when l goes from 6 to 18 (e.g., C6
117 to C6H12), the computation time increase by less than 60% whereas the number of isotopomer is
118 multiplied by $2^{12} = 4096$. This is why the SSA algorithm is well adapted to cross-labeling, e.g.
119 hydrogen carbon, leading to longer chain lengths and thus to a higher combinatoriality.

120 2.3. SSA in ¹³C-DMFA

121 To further test the SSA, we implemented it in a ¹³C-DMFA procedure. A general scheme of
122 the procedure is shown in Figure 4. A series of measurements concerning metabolite concentra-
123 tions and mass isotopomer distributions (MID) with known associated experimental errors is the
124 target of an optimization procedure. The aim is to fit these data with a kinetic model based on
125 mass action laws used for Figure 2 (Table 1). The flux dynamics therefore depend on the kinetic
126 parameters of the reaction laws. Instead of a kinetic model, we could also use the stoichiometric
127 model supplemented with parameterized time functions to describe the flux. The parameter space
128 of the model is then explored to identify the sets of parameters consistent with the target exper-
129 imental data, taking into account the existing uncertainties (see (Valderrama-Bahamóndez and
130 Fröhlich, 2019) for a review of standard method). Once the exploration is completed, the dynam-
131 ics of metabolic fluxes are computed for each selected parameter set, which can be represented as
132 a confidence region for flux trajectories.

133 Here, target datasets were generated for concentrations and mass isotopomers from the DSA
134 at 2, 5, 10, 20, 30, 40s in the same condition as in Figure 3. Then the SSA, with $\Omega = 200$ here,
135 is used to compute the fitness score from target dataset and kinetic parameter set. Two classes
136 of experimental measurements are considered. In a first strategy, only the mass isotopomers $m +$
137 $0 \dots m + 3$ are targeted with an error of 5% (no data provided for the concentrations). In a second
138 strategy, concentration data are also included and targeted with an error of $0.25 \mu\text{M}$. These errors
139 mimic typical experimental and measurement uncertainties. A parameter set is here kept and said
140 to be consistent with target dataset for a chi-square per degree of freedom (i.e., fitness score)
141 remains below unity. Here, a parameter sensitivity analysis computes the parameter ranges, one

Stochastic method for isotope labeling systems

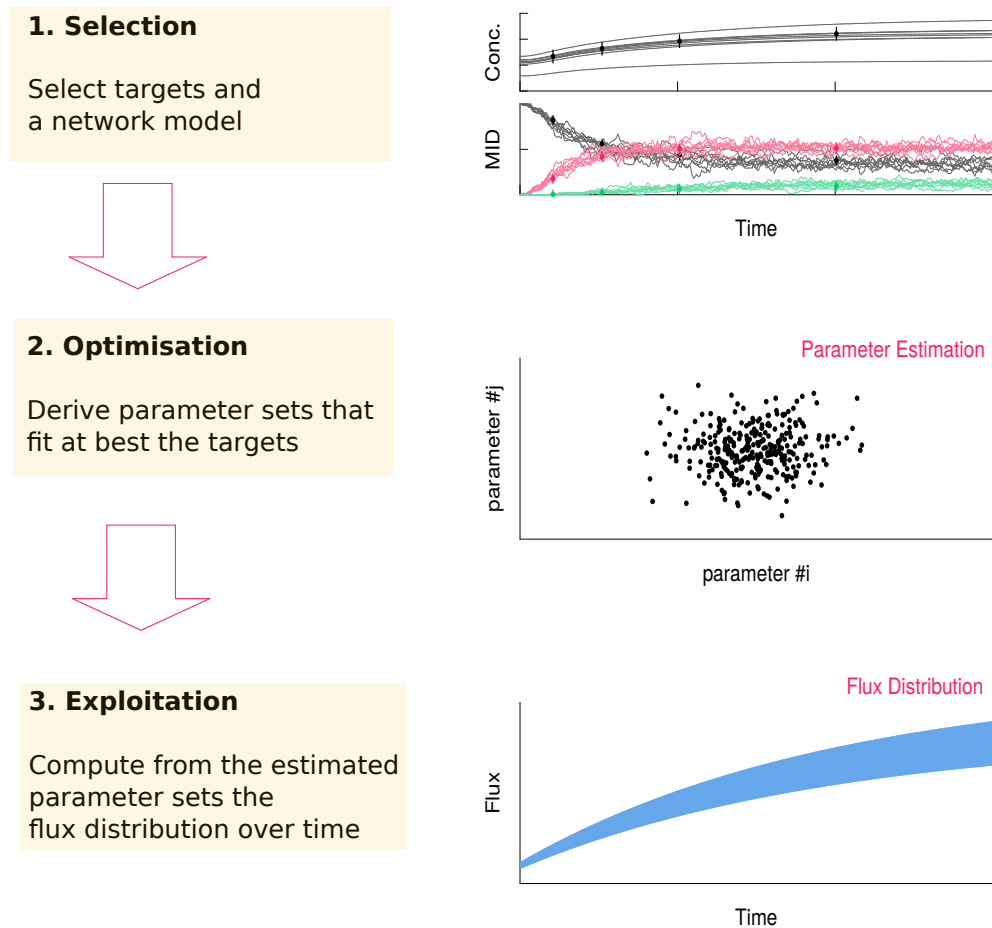


FIG. 4. ^{13}C -DMFA generic workflow using SSA. Metabolite concentration as well as mass isotopomer distribution (MID) are the targets of parameter estimation procedure. Parameter sensitivity analysis provides a list of points in the parameter space that accurately describes targets. The range of flux dynamics is then computed from the parameter value distribution.

142 by one, to illustrate the procedure; the input flux is assumed to be known. For each strategy,
143 the selected parameter sets are finally used to compute the dispersion of the reaction fluxes. As
144 expected, the areas of flux trajectories comprise the exact solution and are reduced when adding
145 concentration data (Figure 5).

146 2.4. SSA in ^{13}C -NMFA or ^{13}C -MFA

147 The stochastic simulation algorithm can also be used to study the propagation of labeling when
148 the network fluxes are in equilibrium. This framework corresponds to ^{13}C -NMFA if the transient

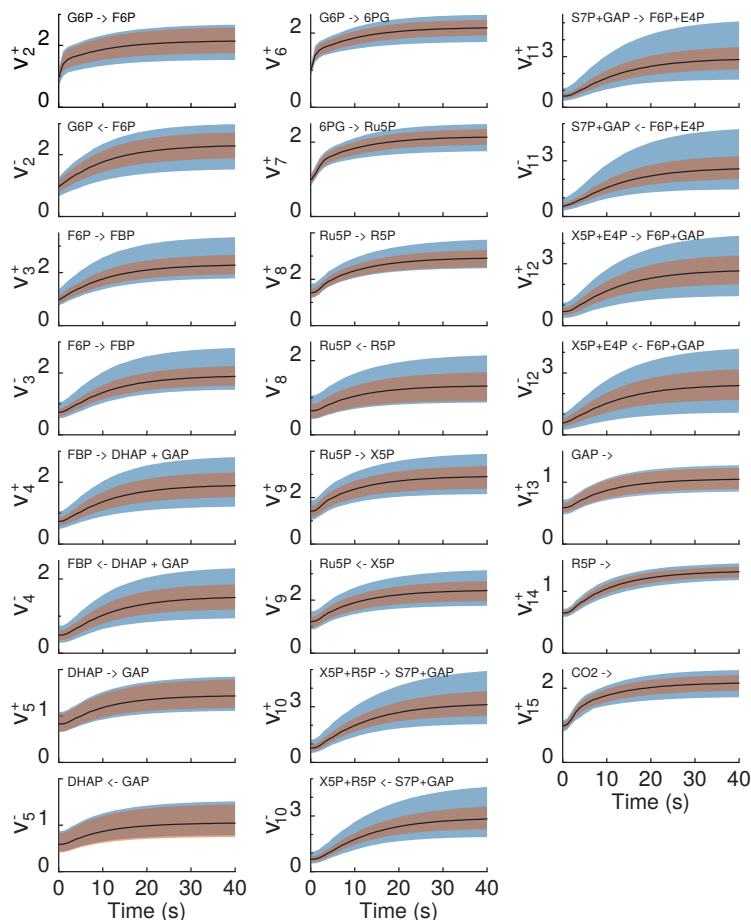


FIG. 5. Directional flux areas compatible with mass isotopomer targeting. The areas are filled with a color code. Blue areas correspond fit strategy 1; red, to fit strategy 2. Solid black lines correspond to the exact solution. $v_i^{+/-}$ correspond to the flux of reaction number i in Table 1 in the forward/barckward direction. The corresponding reaction is also indicated on top of each panel.

149 is studied or to ^{13}C -MFA if the steady state only is studied. For these two frameworks, the state
 150 of the art method uses an EMU decomposition and is implemented in several available software.
 151 Therefore, a comparison has been made between the results and computational times obtained with
 152 the new algorithm presented here on the one hand, and those of the INCA software, implementing
 153 an EMU decomposition Young (2014).

154 Figure 7, which represents the MID dynamics of one of the variables (S7P) obtained with the
 155 two software packages, illustrates the consistency of the results. The calculation of the transient
 156 dynamics was conducted in two networks of different size; on the one hand, the one described by
 157 the reactions of Table 1 (13 species – 15 reactions) and on the other hand, an extended network con-
 158 taining Glycolysis, Pentose Phosphate Pathway, Entner-Doudoroff Pathway, Tricarboxylic Acid

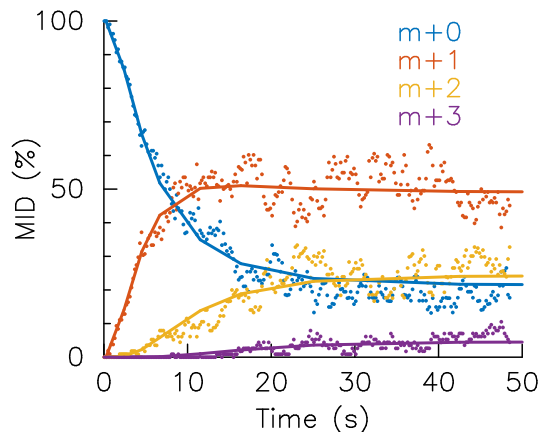


FIG. 6. Comparison of the MIDs dynamics for the S7P variable calculated with the SSA (dots) and INCA software (solid lines). The simulated network corresponds to the reactions listed in Table 1 with equilibrium fluxes (^{13}C -NMFA). The color code is indicated on the figure.

159 Cycle, Amphibolic Reactions, Acetic Acid Formation, and PDO Biosynthesis) summarized in sup-
 160 plementary Table 1 (31 species – 37 reactions). The calculation times for a 50 s integration are
 161 reproduced in Table 2 and show that the SSA used as an forward ^{13}C -NMFA method is signif-
 162 icantly faster than the INCA software. The calculations were performed on the same computer
 163 (Intel(R) Core(TM) i7-10610U at 1.8 GHz) without parallelisation.

164 To complete this comparative study, we implemented the SSA in a ^{13}C -MFA method (only the
 165 steady state is targeted) to determine the reaction fluxes (Figure 7). The larger network (Table S1
 166 of the supplementary material) was used and the MID targets were generated with DSA to know the
 167 exact reaction fluxes. An 2.5% error was used in the fit procedure to mimic experimental errors.
 168 The SSA was coupled to an MCMC method to obtain the accuracy on the determination of the
 169 fluxes corresponding to this error. The INCA software was also used for the flux determination.
 170 The top panel represents the goodness of the fit ($m + 0$, $m + 1$ and $m + 2$ were targeted for each
 171 variable) and the bottom one, the fluxes estimation inferred from the MID targets and their assumed
 172 errors (2.5%). The fluxes distributions derived with MCMC overlap the exact values and INCA
 173 estimations.

	Network 1	Network 2
species	13	31
reactions	15	37
SSA Comput. time	0.025 s	0.08 s
INCA Comput. time	0.5 s	0.5 s

TABLE I. Computational performance of SSA and INCA

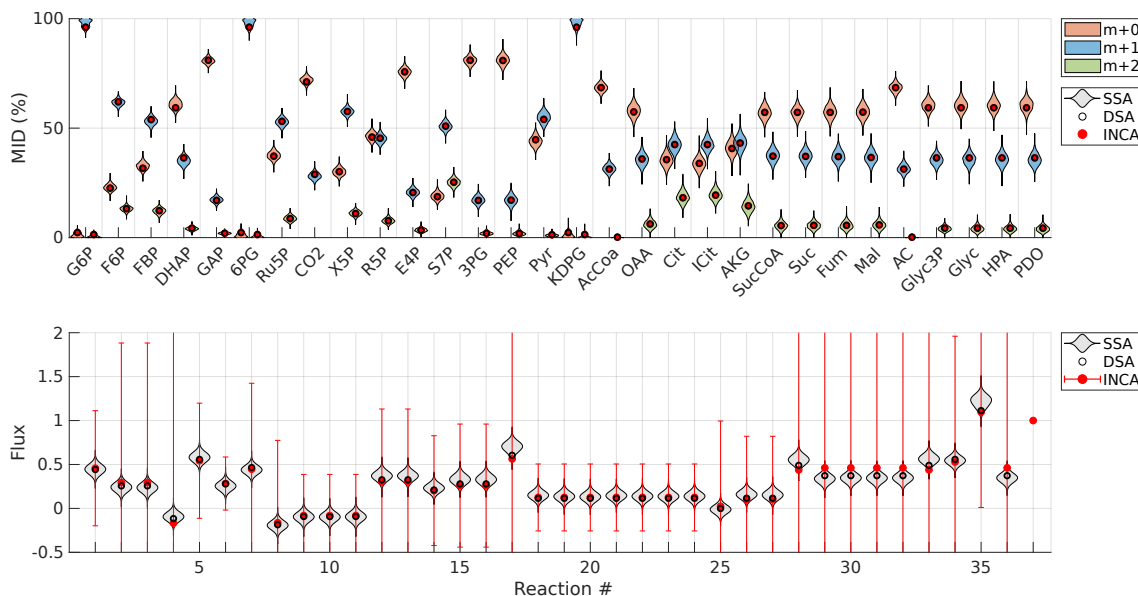


FIG. 7. ^{13}C -MFA with SSA or INCA. A ^{13}C -MFA test performed with SSA coupled to an MCMC method and with INCA on generated MID targets and with a chosen absolute error of 2.5%. Distribution of target fits for each variable (top panel) and estimation of net fluxes (bottom panel). The color code is indicated on the figure.

174 3. THEORY

175 3.1. Chemical master equation model for isotopic labeling networks

176 In a network of (bio)chemical reactions, the temporal evolution of the state probabilities is de-
 177 scribed by the chemical master equation (CME) through a general formalism (Gillespie, 1992).
 178 Deterministic kinetic rate equations, on the one hand, can be derived from the first moments of the
 179 probability distribution and allow for a thorough analysis of the network dynamics by various ana-
 180 lytical techniques (Thompson and Stewart, 2002). The probabilistic features of the dynamics such

181 as bimodal distributions or coefficients of variation, on the other hand, can be investigated with
 182 stochastic implementation of the CME through well-established stochastic simulation algorithms
 183 (Gillespie, 1977, 2001).

184 3.1.1. Isotopomer index and addressing operators

185 A chemical reaction network such as the one depicted in Figure 2 is defined by K reactions
 186 between M chemical species whose concentrations are denoted by S_m with $m \in [1, M]$. The labeling
 187 states of the species S_m is an ordered sequence $(s_{m,1}, s_{m,2}, \dots, s_{m,l_m})$ of length denoted l_m made
 188 of elements $s_{m,i} \in [0, q - 1]$. The species S_m has therefore $L_m = q^{l_m}$ different labeling states or
 189 positional isotopomer indexed by $n_m = \sum_{i=1}^{l_m} s_{m,i} q^{i-1}$, also noted $n_m = (s_{m,1}, s_{m,2}, \dots, s_{m,l_m})_q$ and
 190 called the isotopomer index. A similar approach restricted to $q = 2$ has already been introduced
 191 to describe isotopomer distribution vectors (Schmidt *et al.*, 1997). In the case 13C-labeling, each
 192 carbon may be in two different states (*i.e.*, $q = 2$) and the sequence $(0, 0, 1, 0, 0, 0, 0)$ for S7P
 193 indicates that 13C label is in third carbon position and corresponds to the labeling state number 4,
 194 the S7P species has therefore 2^7 different labeling states.

If the permutation rule is known, one can define addressing operators that compute the iso-
 topomer index of the products from the isotopomer index of the reactants, and *vice versa* for each
 reaction of the network. The addressing operator forms an alternative to atom mapping matrices
 defined by Zupke *et al.* (Zupke and Stephanopoulos, 1994). In the case of the reaction mediated
 by transaldolase (reaction number 11 in Table 1), the addressing operators

$$\sigma_{\text{F6P}}(n_{\text{S7P}}, n_{\text{GAP}}) = (s_{a,1}, s_{a,2}, s_{a,3}, s_{b,1}, s_{b,2}, s_{b,3})_q \quad (1)$$

$$\sigma_{\text{E4P}}(n_{\text{S7P}}, n_{\text{GAP}}) = (s_{a,4}, s_{a,5}, s_{a,6}, s_{a,7})_q, \quad (2)$$

compute the product index from reactant index $n_{\text{S7P}} = (s_{a,1}, \dots, s_{a,7})_q$ and $n_{\text{GAP}} = (s_{b,1}, s_{b,2}, s_{b,3})_q$.

In the same manner,

$$\sigma_{\text{S7P}}(n_{\text{F6P}}, n_{\text{E4P}}) = (s_{c,1}, s_{c,2}, s_{c,3}, s_{d,1}, s_{d,2}, s_{d,3}, s_{d,4})_q \quad (3)$$

$$\sigma_{\text{GAP}}(n_{\text{F6P}}, n_{\text{E4P}}) = (s_{c,4}, s_{c,5}, s_{c,6})_q \quad (4)$$

195 compute the reactant index from product index $n_{\text{F6P}} = (s_{c,1}, \dots, s_{c,6})_q$ and $n_{\text{E4P}} = (s_{d,1}, \dots, s_{d,4})_q$.
 196 Therefore, in the context of a 13C labeling ($q = 2$), the reaction between a doubly labeled S7P
 197 (1000100) and a simply labeled GAP (001) – *i.e.* $n_{\text{S7P}} = 17$ and $n_{\text{GAP}} = 4$ – produces an F6P
 198 (100001) and an E4P (0100) – *i.e.* $n_{\text{F6P}} = 33$ and $n_{\text{E4P}} = 2$.

199 **3.1.2. Chemical master equation description**

200 The chemical master equation (CME) is a general and accurate formalism to describe the
 201 stochastic dynamics in (bio)chemical reaction networks (Gillespie, 2000). This formalism can
 202 be easily extended to also describe the stochastic dynamics of labeling states of chemical species.
 203 With above notations for isotopomer index and addressing operators, the probabilistic dynamics in
 204 isotope labeling network can naturally be described in the CME framework. The chemical species
 205 with the largest number of isotopomers determines the size N of the state space ($N = 2^7$ in the ex-
 206 ample used in Figure 2). The state of the whole system is therefore described by a $M \times N$ integer
 207 matrix ω : $\omega_{m,n}$ indicating the number of m th species in the n th labeling state. The total number of
 208 the m th species is noted $\Omega S_m = \sum_n \omega_{m,n}$ where Ω is a volume (involved as a scaling factor) and S_m
 209 is a concentration. The probability that the internal sequence of the m th species corresponds to the
 210 n th isotopomer is denoted $\rho_{m,n} = \omega_{m,n} / \Omega S_m$. The formalism of the CME describes the temporal
 211 evolution of the probability of the system to be in the state ω , noted $\mathcal{P}_\omega(t)$.

212 The chemical reactions that define the network are characterized by both a concentration-
 213 dependent flux of reagents v_k with $k \in [0, K]$ and a permutation rule between the position of labeled
 214 atoms of reactants and products. Reactions are distinguished depending on their input, output, or
 215 internal position in the network. For instance, the network depicted in Figure 2 has one input reac-
 216 tion, 3 output reactions and 11 internal reactions. As seen latter on, input reactions always require
 217 a particular consideration since the reactant is not modified. For keeping notations simple, we
 218 restrict to Bi Bi reactions of the type $A + B \rightarrow C + D$ where A, B, C, D are either chemical species
 219 or empty sets. In this case, the CME reads,

$$\begin{aligned} \frac{d}{dt} \mathcal{P}_\omega(t) = & \sum_{k=1}^{K_n} \Omega v_k(S) \sum_{n,n'} \left[\mathbb{E}_{a_k,n}^+ \mathbb{E}_{b_k,n'}^+ \mathbb{E}_{c_k,\sigma_{c_k}(n,n')}^- \mathbb{E}_{d_k,\sigma_{d_k}(n,n')}^- - 1 \right] \rho_{a_k,n} \rho_{b_k,n'} \mathcal{P}_\omega(t) \quad (5) \\ & + \sum_{k=K_n+1}^{K_n+K_i} \Omega v_k(S) \sum_n I_{c_k,n}(t) \left[\mathbb{E}_{c_k,n}^- - 1 \right] \mathcal{P}_\omega(t) \end{aligned}$$

220 The integers a_k, b_k, c_k, d_k correspond to the indices of the species A, B, C, D of the k th reaction of
 221 type $A + B \rightarrow C + D$, the integer being null in the case of an empty set. The writing uses scale
 222 operators $\mathbb{E}_{m,n}^\pm$ (Van Kampen, 1992) :

$$\mathbb{E}_{m,n}^\pm \rho_{m_1,n_1} \rho_{m_2,n_2} \mathcal{P}_\omega(t) = \left[\rho_{m_1,n_1} \pm \frac{\delta_{m,m_1} \delta_{n,n_1}}{\Omega S_{m_1}} \right] \left[\rho_{m_2,n_2} \pm \frac{\delta_{m,m_2} \delta_{n,n_2}}{\Omega S_{m_2}} \right] \mathcal{P}_{\omega \pm \mathbf{E}_{m,n}}(t) \quad (6)$$

223 where $\mathbf{E}_{m,n}$ is a matrix of the canonical base (only the element at the intersection of row m and
 224 column n is non-zero and is unity), and $\delta_{i,j}$ the Kronecker symbol (unity if indexes are equal,
 225 zero either). The K_n first reactions concern internal and output reactions whereas the remaining
 226 K_I concerns input reactions ($\emptyset \rightarrow C$). In this later case, $I_{c_k,n}$ is the fixed probability to have an
 227 isotopomer n of the input species c_k .

228 3.2. Derived Deterministic Simulation Algorithm (DSA)

The CME (Eq. 5) can be approximated in the large size limit $\Omega \rightarrow \infty$ by deterministic rate
 equation dynamics. The probability of each internal sequence is denoted by $\rho_{m,n}$ such that $S_{m,n} =$
 $S_m \rho_{m,n}$ describes the concentration of n th-isotopomer of the m th species. The time evolution of
 isotopomer concentrations is governed by the distribution rules specific to each reaction and is
 formalized mathematically by a permutation of the concatenated internal sequence between the
 reagents and the products. The deterministic system evolves according to the ordinary differential
 equations,

$$\frac{d}{dt} S_{m,n} = \sum_{k=1}^K N_{m,k} v_k(S) \Phi_{m,n}^k(\rho), \quad (7)$$

229 where N denotes the stoichiometry matrix, v_k $k \in [0, K]$ the concentration-dependent flux of
 230 reagents, and $\Phi_{m,n}^k(\rho)$ the flux fraction describing the permutation rules of chemical reaction sat-
 231 isfying $\sum_n \Phi_{m,n}^k(\rho) = 1$. Algorithm to simulate Eq. 7 with standard Runge-Kutta-Fehlberg method
 232 of order 5 with adaptive step is called Deterministic Simulation Algorithm (DSA).

Let us first consider the internal reactions restricted to Bi Bi reactions of the form $A + B \rightarrow$
 $C + D$. The reaction is characterized by the reordering of atom position defining addressing oper-
 ations of the products according to the indices of the reagents. The operator $\sigma_A(n_c, n_d)$ gives the
 isotopomer index of the A species that produce C and D of isotopomer index n_c and n_d , respec-
 tively. In that case, the reaction index k is omitted and the flux fraction reads,

$$\Phi_{a,n_a}(\rho) = \rho_{a,n_a} \quad (8)$$

$$\Phi_{b,n_b}(\rho) = \rho_{b,n_b} \quad (9)$$

$$\Phi_{c,n_c}(\rho) = \sum_{n_d=0}^{L_d-1} \rho_{a,\sigma_A(n_c,n_d)} \rho_{b,\sigma_B(n_c,n_d)} \quad (10)$$

$$\Phi_{d,n_d}(\rho) = \sum_{n_c=0}^{L_c-1} \rho_{a,\sigma_A(n_c,n_d)} \rho_{b,\sigma_B(n_c,n_d)} \quad (11)$$

233 where L_c and L_d are the number of isotopomers of C and D species. If B is an empty set then
 234 $\rho_{b,x} = 1$, and if D is an empty set Eq. (11) is useless. If the reaction is an output reaction, then
 235 C and D are empty sets, $\Phi_{a,n_a}(\rho)$ and $\Phi_{b,n_b}(\rho)$ are computed with the same rules as internal
 236 reactions. As mentioned, the input reactions must be treated separately since the reactants are not
 237 variables but parameters. We consider here input reactions of simple form $\emptyset \rightarrow C$ and we note I_n
 238 the probability of synthesis of the species C in the state n , thus $\Phi_{c,n_c}(\rho) = I_n$ in this case.

Alternatively, the dynamical system (Eq. 7) can also be written as

$$\frac{d}{dt} S_m = \sum_{k=1}^K N_{m,k} v_k(S) \quad m \in [0, M] \quad (12)$$

$$S_m \frac{d}{dt} \rho_{m,n} = \sum_{k=1}^K N_{m,k} v_k(S) \left(\Phi_{m,n}^k(\rho) - \rho_{m,n} \right) \quad (13)$$

239 The first equation describes the time evolution of species concentrations while the second equation
 240 describes the time evolution of the fraction of different isotopomers. This additional equation high-
 241 lights the key role of the concentrations S_m in the timescale of changes in isotopomer distribution:
 242 higher concentration values lead to slower evolution of isotopomer distributions.

243 The permutation rules defined in Φ may be easily extended to more complex reactions. In the
 244 framework developed here, they only depend on the permutation relations and not on the mathe-
 245 matical forms of concentration-dependent flux, because we assume that the internal modification
 246 does not impact the reaction rate. If the construction rules are simple to establish and to implement
 247 in a numerical code, the computation time of the flux vector of the dynamic system (the right-hand
 248 side term of Eq. 7) increases significantly with the length of the sequences and the number of
 249 isotopomers, because of the many summations of terms. Moreover, this implementation computes
 250 the evolution of all possible isotopomers while the experimental labeling used nowadays gener-
 251 ates only a small subset of the possible isotopomers (Metallo, Walther, and Stephanopoulos, 2009).
 252 The deterministic system therefore requires a large number of unnecessary calculations even with
 253 an optimized implementation. It nevertheless serves as a useful benchmark to check the relative
 254 accuracy and efficiency of other methods.

255 3.3. Derived Stochastic Simulation Algorithm (SSA)

256 The CME is in fact a continuous-time approximation of discrete time stochastic processes.
 257 Stochastic algorithms are often used to simulate the molecular dynamics in chemical reaction

258 networks and capture the statistical and temporal features of fluctuations. In the case of the above
 259 CME (Eq. 5), time evolution of isotopomer distribution can also be simulated by a stochastic
 260 Monte-Carlo algorithm based on the next reaction methods (Gibson and Bruck, 2000), here called
 261 Stochastic Simulation Algorithm (SSA). Each chemical species is represented by a finite sample
 262 of isotopomers (Figure 2) where the sampling size is proportional to the concentration of the
 263 corresponding chemical species.

264 The sample size of the variable m is ΩS_m where Ω is a volume. The occurrence of a chem-
 265 ical reaction is determined by the standard next reaction methods that we have adapted, SSA is
 266 summarized as:

267 Init: Compute the reaction time for all reactions

$$\tau_k = \frac{1}{\Omega v_k(S)} \quad (14)$$

268 Step 1: Find the smallest reaction time $\tau_{k'} = \min(\tau_k)$ and do reaction k' by randomly picking the
 269 reagents from their samples and synthesizing the products following the permutation rule of
 270 the reaction;

271 Step 2: Increment time t by $\tau_{k'}$ and compute a next time for reaction k ;

272 Step 3: Adjust the set of reaction times to account for sample size variation induced by reaction k'

$$\tau_k \leftarrow \frac{v_{k,\text{old}}}{v_{k,\text{new}}} (\tau_k - t) + t$$

273 and iterate to Step 1

274 In this sequential process, each stochastic occurrence of a chemical reaction induces discrete
 275 changes in the number of species and of isotopomers associated to each chemical species, which
 276 results in stochastic fluctuations of both species concentrations and isotopomer distribution.

277 Contrary to the common use of stochastic simulation algorithms for chemical reaction net-
 278 works, the Ω value does not have to represent the real number of molecules for a reference con-
 279 centration because the algorithm considers mainly the propagation of marked atoms and not the
 280 stochastic fluctuations of the chemical reactions linked to the finite number of copies. In the con-
 281 text of metabolic networks, fluctuation of the reaction times τ_k are indeed rarely relevant. Because
 282 of the high copy number of metabolites, numerous reactions occur and fluctuations of the reaction
 283 times τ_k do not induce much concentration fluctuations. If, however, one wished to decline this

284 algorithm to study the stochastic fluctuation of the chemical reactions, it would be enough to use
285 the relation $\tau_k = \frac{1}{\Omega v_k(S)} \log\left(\frac{1}{U_k}\right)$ with independent uniform random deviates U_k in $[0, 1]$ for the
286 reaction time computations. In the latter case, a realistic estimate for the Ω parameter value must
287 be used.

288 4. METHODS

289 4.1. Code availability and computer simulation

290 Both methods, SSA and DSA, were implemented with the same highest level of optimization,
291 using low-level bit-manipulation tools to implement addressing operators in a Fortran code (com-
292 piled with gfortran and optimization flag “-O3”). Simulations were run on a standard laptop with
293 an Intel(R) Core(TM) i5-6300U CPU at 2.40GHz. No parallelization were used. The fortran code
294 is available in github
295 <https://github.com/Qthommen/Stochastic-method-for-isotope-labeling-systems>.
296 git

297 4.2. Goodness of Fit

298 The chi-square per degree of freedom $\chi_v^2 = \frac{1}{n-p} \sum_{i=1}^n \frac{(y_i - y_i^*)^2}{\sigma_i^2}$ is used a goodness of fit criterion.
299 n is the number of targets; p , the number of parameters; y_i et y_i^* the computations and targets; σ_i^2
300 the variance. The fit is accepted $\chi_v^2 < 1$.

301 4.3. Metabolic Network

302 Table 1 lists the chemical reactions and carbon rearrangements of the upper part of the increased
303 glycolysis of the pentose phosphate pathway (Figure 2). To illustrate the dynamics of the propa-
304 gation of the labeled carbons and for the sake of simplicity, the reaction rate used corresponds to
305 the mass action law with a kinetic parameter of unit value.

reac nb	chemical reaction	reaction speed	parameter values
1	$\text{GLU (abcde)} \longrightarrow \text{G6P (abcde)}$	$v_{G6P} = k_1$	$k_1 = 1 \rightarrow 2 \mu\text{M/s}$
2	$\text{G6P (abcdef)} \rightleftharpoons \text{F6P (abcdef)}$	$v_2 = k_2^+ [\text{G6P}] - k_2^- [\text{F6P}]$	$k_2^+ = k_2^- = 1/\text{s}$
3	$\text{F6P (abcdef)} \rightleftharpoons \text{FBP (abcdef)}$	$v_3 = k_3^+ [\text{F6P}] - k_3^- [\text{FBP}]$	$k_3^+ = k_3^- = 1/\text{s}$
4	$\text{FBP (acbddef)} \rightleftharpoons \text{DHAP (cba) + GAP (def)}$	$v_4 = k_4^+ [\text{FBP}] - k_4^- [\text{DHAP}][\text{GAP}]$	$k_4^+ = 1/\text{s}; k_4^- = 1/\mu\text{M/s}$
5	$\text{DHAP (abc)} \rightleftharpoons \text{GAP (abc)}$	$v_5 = k_5^+ [\text{DHAP}] - k_5^- [\text{GAP}]$	$k_5^+ = k_5^- = 1/\text{s}$
6	$\text{G6P (abcdef)} \longrightarrow \text{6PG (abcdef)}$	$v_6 = k_6 [\text{G6P}]$	$k_6 = 1/\text{s}$
7	$\text{6PG (abcdef)} \longrightarrow \text{CO2 (a) + Ru5 (bcdef)}$	$v_7 = k_7 [\text{6PG}]$	$k_7 = 1/\text{s}$
8	$\text{Ru5 (abcde)} \rightleftharpoons \text{R5P (abcde)}$	$v_8 = k_8^+ [\text{Ru5}] - k_8^- [\text{R5P}]$	$k_8^+ = k_8^- = 1/\text{s}$
9	$\text{Ru5 (abcde)} \rightleftharpoons \text{X5P (abcde)}$	$v_9 = k_9^+ [\text{Ru5}] - k_9^- [\text{X5P}]$	$k_9^+ = k_9^- = 1/\text{s}$
10	$\text{X5P (abcde) + R5P (ABCDE)} \rightleftharpoons \text{S7P (abABCDE) + GAP (cde)}$	$v_{10} = k_{10}^+ [\text{X5P}][\text{R5P}] - k_{10}^- [\text{S7P}][\text{GAP}]$	$k_{10}^+ = k_{10}^- = 1/\mu\text{M/s}$
11	$\text{S7P (abcdefg) + GAP (ABC)} \rightleftharpoons \text{F6P (abcABC) + E4P (defg)}$	$v_{11} = k_{11}^+ [\text{S7P}][\text{GAP}] - k_{11}^- [\text{F6P}][\text{E4P}]$	$k_{11}^+ = k_{11}^- = 1/\mu\text{M/s}$
12	$\text{X5P (abcde) + E4P (ABCD)} \rightleftharpoons \text{F6P (abABCD) + GAP (cde)}$	$v_{12} = k_{12}^+ [\text{X5P}][\text{E4P}] - k_{12}^- [\text{F6P}][\text{GAP}]$	$k_{12}^+ = k_{12}^- = 1/\mu\text{M/s}$
13	$\text{GAP (abc)} \longrightarrow$	$v_{13} = k_{13} [\text{GAP}]$	$k_{13} = 1/\text{s}$
14	$\text{R5P (abcde)} \longrightarrow$	$v_{14} = k_{14} [\text{R5P}]$	$k_{14} = 1/\text{s}$
15	$\text{CO2 (a)} \longrightarrow$	$v_{15} = k_{15} [\text{CO2}]$	$k_{15} = 1/\text{s}$

TABLE II. List of chemical reactions and reaction speed used to simulate the ^{13}C propagation through the upper glycolytic pathways supplemented by the pentose phosphate pathway displayed in Figure 2

5. DISCUSSION

In this study, we propose a stochastic algorithm to emulate the propagation of labeled atoms in a nonstationary metabolic system. This algorithm derives from the chemical master equation which is the most comprehensive framework for describing chemical reaction network dynamics. The efficiency of the algorithm has been applied to ^{13}C -DMFA of the illustrative case of the pentose phosphate pathways for which ^{13}C -labeling and concentration time series data has been synthesized. One of the main computational advantages of the proposed method lies in the very weak dependence of the computation time on the length of the marking chain and thus the number of isotopomer. Deterministic methods exhibit by construction a number of variables and a computation time that both rapidly increase with the combinatoriality associated to the power-law dependence with the chain length. SSA is therefore well adapted to the study of parallel labeling, combining for instance carbon and hydrogen labeling (Lewis *et al.*, 2014; Antoniewicz, 2015b; Jacobson *et al.*, 2019; Dong *et al.*, 2019). Moreover, a simulation that mimics the stochastic and discrete nature of metabolic reaction processes provides a more accurate and comprehensive picture relating the propagation of labeling with the dynamics of isotopomer and metabolite concentrations. Finally, this rigorous and straightforward method requires no tinkering or approximations depending on the resolution of the experimental measurements or the nature of the metabolic process, as it calculates all isotopomers at no additional cost and natively handles both stationary and non-stationary conditions. In other words, the SSA method can be used interchangeably or simultaneously for ^{13}C -MFA (Hurbain *et al.*, 2022), ^{13}C -NMFA or ^{13}C -DMFA. Because problem-dependent reduction or solving techniques are not used, the implementation does not require any particular software and can simply be done in any programming language, as it is the case for chemical kinetics modeling (see Code availability).

Estimation of metabolic flux dynamics from ^{13}C labeling and metabolomics data can be done either by inverse kinetic model modeling (Wahl, Nöh, and Wiechert, 2008; Baxter *et al.*, 2007) or by considering flux function (Antoniewicz *et al.*, 2007; Leighty and Antoniewicz, 2011; Schumacher and Wahl, 2015; Quek *et al.*, 2020). The preference of latter methods have been motivated by the lack of information about intracellular enzyme kinetics, but also the computational cost of deterministic simulation of kinetic models comprising isotopomer variables. Thanks to the computational efficiency of SSA for simulating isotopomer dynamics, inverse kinetic modeling integrating ^{13}C labeling data become an achievable goal. However, SSA can still be used

337 with flux function as well, for instance with constant function in case of stationary metabolic
338 condition (*e.g.*, ^{13}C -MFA and ^{13}C -NMFA) (Hurbain *et al.*, 2022). To summarize, a SSA-based
339 ^{13}C -DMFA method would require (1) defining a stoichiometry model, (2) defining kinetic laws or
340 flux functions, (3) using an optimization method to estimate the parameters of reaction laws or flux
341 functions, (4) using a Monte-Carlo method to evaluate the distribution of such parameters, (5) ad-
342 just iteratively the model (stoichiometry or kinetic structure) to optimize tradeoff between a good
343 fit and a narrow parameter distributions. The last step corresponds to the well-known problem of
344 model selection (Mangan *et al.*, 2017). It is, however, to keep in mind that overparametrization is
345 not a issue as long as one focuses on the estimation of flux trajectories. If, on the other hand, the
346 ^{13}C -DMFA is used for dynamic control purposes (Hartline *et al.*, 2021), the parameterization of
347 flux functions will be of great importance, and it will be necessary to model the chemical kinetics
348 as precisely as possible.

349 The only delicate issue associated to this method is associated to the appropriate choice of the
350 sample size Ω . Ω must be large enough to ensure that the level of fluctuations in isotopomer
351 concentration are below the experimental uncertainties. At the same time, computational time
352 scales linearly with Ω motivating to keep its value as low as possible. The parameter Ω thus needs
353 to be adjusted to a typical value (typically 100 – 1000) to optimize the tradeoff between simulation
354 uncertainties (below experimental uncertainties) and computational efficiency. For such system
355 size, the residual fluctuations of isotopomer concentration leads to a narrow distribution of error
356 score for a same parameter set, which is not an issue when using Monte Carlo sampling algorithm
357 used for metabolic flux analysis (Theorell *et al.*, 2017; Valderrama-Bahamóndez and Fröhlich,
358 2019; Heinonen *et al.*, 2019; Theorell and Nöh, 2020). For a given value Ω , a temporal averaging
359 procedure may be added to narrow the distribution of mass isotopomer concentrations for given
360 Ω , allowing to use lower Ω values. Another limitation relates to the high number of reactions
361 which depends on the absolute value of directional fluxes, not of net fluxes. This limitation can be
362 largely compensated by the property that the number of operations (*e.g.*, computational time) does
363 not depend on isotopomer number per metabolites.

365 **ACKNOWLEDGEMENTS**

366 The authors thank Darka Labavic for the fruitful discussions. Canther Laboratory is part of
367 ONCOLille institute.

368 **FUNDING**

369 This work has been supported by the LABEX CEMPI (ANR-11-LABX-0007) and by the Min-
370 istry of Higher Education and Research, Hauts de France council and European Regional Devel-
371 opment Fund (ERDF) through the Contrat de Projets Etat-Region (CPER Photonics for Society
372 P4S and CPER Cancer 2015-2020).

373 **REFERENCES**

- 374 Allen, D. K. and Young, J. D., “Tracing metabolic flux through time and space with isotope label-
375 ing experiments,” *Current opinion in biotechnology* **64**, 92–100 (2020).
- 376 Antoniewicz, M. R., “Methods and advances in metabolic flux analysis: a mini-review,” *Journal*
377 *of industrial microbiology and biotechnology* **42**, 317–325 (2015a).
- 378 Antoniewicz, M. R., “Parallel labeling experiments for pathway elucidation and ¹³C metabolic
379 flux analysis,” *Current opinion in biotechnology* **36**, 91–97 (2015b).
- 380 Antoniewicz, M. R., “A guide to metabolic flux analysis in metabolic engineering: methods, tools
381 and applications,” *Metabolic engineering* **63**, 2–12 (2021).
- 382 Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G., “Elementary metabolite units
383 (EMU): a novel framework for modeling isotopic distributions,” *Metabolic engineering* **9**, 68–86
384 (2007).
- 385 Antoniewicz, M. R., Kraynie, D. F., Laffend, L. A., González-Lergier, J., Kelleher, J. K., and
386 Stephanopoulos, G., “Metabolic flux analysis in a nonstationary system: fed-batch fermentation
387 of a high yielding strain of *E. coli* producing 1, 3-propanediol,” *Metabolic engineering* **9**, 277–
388 292 (2007).
- 389 Baxter, C., Liu, J., Fernie, A., and Sweetlove, L., “Determination of metabolic fluxes in a non-
390 steady-state system,” *Phytochemistry* **68**, 2313–2319 (2007).

- 391 Bouzier-Sore, A.-K. and Bolaños, J. P., “Uncertainties in pentose-phosphate pathway flux assess-
392 ment underestimate its contribution to neuronal glucose consumption: relevance for neurode-
393 generation and aging,” *Frontiers in Aging Neuroscience* **7**, 89 (2015).
- 394 Creek, D. J., Mazet, M., Achcar, F., Anderson, J., Kim, D.-H., Kamour, R., Morand, P., Millerioux,
395 Y., Biran, M., Kerkhoven, E. J., *et al.*, “Probing the metabolic network in bloodstream-form
396 trypanosoma brucei using untargeted metabolomics with stable isotope labelled glucose,” *PLoS*
397 *pathogens* **11**, e1004689 (2015).
- 398 Crown, S. B. and Antoniewicz, M. R., “Parallel labeling experiments and metabolic flux analysis:
399 Past, present and future methodologies,” *Metabolic engineering* **16**, 21–32 (2013).
- 400 Diaz-Moralli, S., Aguilar, E., Marin, S., Coy, J. F., Dewerchin, M., Antoniewicz, M. R., Meca-
401 Cortés, O., Notebaert, L., Ghesquière, B., Eelen, G., *et al.*, “A key role for transketolase-like 1
402 in tumor metabolic reprogramming,” *Oncotarget* **7**, 51875 (2016).
- 403 Dong, W., Moon, S. J., Kelleher, J. K., and Stephanopoulos, G., “Dissecting mammalian cell
404 metabolism through ¹³C-and ²H-isotope tracing: Interpretations at the molecular and systems
405 levels,” *Industrial & Engineering Chemistry Research* **59**, 2593–2610 (2019).
- 406 Gibson, M. A. and Bruck, J., “Efficient exact stochastic simulation of chemical systems with many
407 species and many channels,” *The journal of physical chemistry A* **104**, 1876–1889 (2000).
- 408 Gillespie, D. T., “Exact stochastic simulation of coupled chemical reactions,” *The journal of phys-
409 ical chemistry* **81**, 2340–2361 (1977).
- 410 Gillespie, D. T., “A rigorous derivation of the chemical master equation,” *Physica A: Statistical
411 Mechanics and its Applications* **188**, 404–425 (1992).
- 412 Gillespie, D. T., “The chemical langevin equation,” *The Journal of Chemical Physics* **113**, 297–306
413 (2000).
- 414 Gillespie, D. T., “Approximate accelerated stochastic simulation of chemically reacting systems,”
415 *The Journal of chemical physics* **115**, 1716–1733 (2001).
- 416 Hartline, C. J., Schmitz, A. C., Han, Y., and Zhang, F., “Dynamic control in metabolic engineer-
417 ing: Theories, tools, and applications,” *Metabolic engineering* **63**, 126–140 (2021).
- 418 Heinonen, M., Osmala, M., Mannerström, H., Wallenius, J., Kaski, S., Rousu, J., and Lähdesmäki,
419 H., “Bayesian metabolic flux analysis reveals intracellular flux couplings,” *Bioinformatics* **35**,
420 i548–i557 (2019).
- 421 Hurbain, J., Thommen, Q., Anquez, F., and Pfeuty, B., “Quantitative modeling of pentose phos-
422 phate pathway response to oxidative stress reveals a cooperative regulatory strategy,” *bioRxiv*

423 (2022), 10.1101/2022.02.04.478659.

424 Jacobson, T. B., Adamczyk, P. A., Stevenson, D. M., Regner, M., Ralph, J., Reed, J. L., and
425 Amador-Noguez, D., “²H and ¹³C metabolic flux analysis elucidates in vivo thermodynamics
426 of the ed pathway in *zymomonas mobilis*,” *Metabolic engineering* **54**, 301–316 (2019).

427 Kuehne, A., Emmert, H., Soehle, J., Winnefeld, M., Fischer, F., Wenck, H., Gallinat, S., Terstegen,
428 L., Lucius, R., Hildebrand, J., *et al.*, “Acute activation of oxidative pentose phosphate pathway
429 as first-line response to oxidative stress in human skin cells,” *Molecular cell* **59**, 359–371 (2015).

430 Lee, M. H., Malloy, C. R., Corbin, I. R., Li, J., and Jin, E. S., “Assessing the pentose phosphate
431 pathway using [2, 3-¹³C₂] glucose,” *NMR in Biomedicine* **32**, e4096 (2019).

432 Leighty, R. W. and Antoniewicz, M. R., “Dynamic metabolic flux analysis (DMFA): a frame-
433 work for determining fluxes at metabolic non-steady state,” *Metabolic engineering* **13**, 745–755
434 (2011).

435 Lewis, C. A., Parker, S. J., Fiske, B. P., McCloskey, D., Gui, D. Y., Green, C. R., Vokes, N. I.,
436 Feist, A. M., Vander Heiden, M. G., and Metallo, C. M., “Tracing compartmentalized nadph
437 metabolism in the cytosol and mitochondria of mammalian cells,” *Molecular cell* **55**, 253–263
438 (2014).

439 Luo, H., Shen, T., and Xie, X., “Stochastic simulation of enzymatic kinetics for ¹³C isotope
440 labeling at the single-cell scale,” *Reaction Kinetics, Mechanisms and Catalysis* **135**, 2341–2355
441 (2022).

442 Mangan, N. M., Kutz, J. N., Brunton, S. L., and Proctor, J. L., “Model selection for dynamical
443 systems via sparse regression and information criteria,” *Proceedings of the Royal Society A:*
444 *Mathematical, Physical and Engineering Sciences* **473**, 20170009 (2017).

445 Metallo, C. M., Walther, J. L., and Stephanopoulos, G., “Evaluation of ¹³C isotopic tracers for
446 metabolic flux analysis in mammalian cells,” *Journal of biotechnology* **144**, 167–174 (2009).

447 Niedenführ, S., Wiechert, W., and Nöh, K., “How to measure metabolic fluxes: a taxonomic guide
448 for ¹³C fluxomics,” *Current opinion in biotechnology* **34**, 82–90 (2015).

449 Ohno, S., Quek, L.-E., Krycer, J. R., Yugi, K., Hirayama, A., Ikeda, S., Shoji, F., Suzuki, K.,
450 Soga, T., James, D. E., *et al.*, “Kinetic trans-omic analysis reveals key regulatory mechanisms
451 for insulin-regulated glucose metabolism in adipocytes,” *Iscience* **23**, 101479 (2020).

452 Quek, L.-E., Krycer, J. R., Ohno, S., Yugi, K., Fazakerley, D. J., Scalzo, R., Elkington, S. D., Dai,
453 Z., Hirayama, A., Ikeda, S., *et al.*, “Dynamic ¹³C flux analysis captures the reorganization of
454 adipocyte glucose metabolism in response to insulin,” *Iscience* **23**, 100855 (2020).

Stochastic method for isotope labeling systems

- 455 Schmidt, K., Carlsen, M., Nielsen, J., and Villadsen, J., “Modeling isotopomer distributions in
456 biochemical networks using isotopomer mapping matrices,” *Biotechnology and bioengineering*
457 **55**, 831–840 (1997).
- 458 Schumacher, R. and Wahl, S. A., “Effective estimation of dynamic metabolic fluxes using ¹³C
459 labeling and piecewise affine approximation: from theory to practical applicability,” *Metabolites*
460 **5**, 697–719 (2015).
- 461 Selivanov, V. A., Puigjaner, J., Sillero, A., Centelles, J. J., Ramos-Montoya, A., Lee, P. W.-N.,
462 and Cascante, M., “An optimized algorithm for flux estimation from isotopomer distribution in
463 glucose metabolites,” *Bioinformatics* **20**, 3387–3397 (2004).
- 464 Stephanopoulos, G., “Metabolic fluxes and metabolic engineering,” *Metabolic engineering* **1**, 1–11
465 (1999).
- 466 Theorell, A., Leweke, S., Wiechert, W., and Nöh, K., “To be certain about the uncertainty:
467 Bayesian statistics for ¹³C metabolic flux analysis,” *Biotechnology and bioengineering* **114**,
468 2668–2684 (2017).
- 469 Theorell, A. and Nöh, K., “Reversible jump MCMC for multi-model inference in metabolic flux
470 analysis,” *Bioinformatics* **36**, 232–240 (2020).
- 471 Thompson, J. M. T. and Stewart, H. B., *Nonlinear dynamics and chaos* (John Wiley & Sons, 2002).
- 472 Valderrama-Bahamóndez, G. I. and Fröhlich, H., “Mcmc techniques for parameter estimation of
473 ode based models in systems biology,” *Frontiers in Applied Mathematics and Statistics* **5**, 55
474 (2019).
- 475 Van Kampen, N. G., *Stochastic processes in physics and chemistry*, Vol. 1 (Elsevier, 1992).
- 476 Wahl, S. A., Nöh, K., and Wiechert, W., “¹³C labeling experiments at metabolic nonstationary
477 conditions: an exploratory study,” *Bmc Bioinformatics* **9**, 1–18 (2008).
- 478 Young, J. D., “INCA: a computational platform for isotopically non-stationary metabolic flux
479 analysis,” *Bioinformatics* **30**, 1333–1335 (2014).
- 480 Zupke, C. and Stephanopoulos, G., “Modeling of isotope distributions and intracellular fluxes
481 in metabolic networks using atom mapping matrixes,” *Biotechnology Progress* **10**, 489–498
482 (1994).