Stochastic simulation algorithm for isotope-based dynamic flux analysis

Quentin Thommen, Julien Hurbain, Benjamin Pfeuty

Stochastic simulation algorithm for isotope-based dynamic flux analysis

Quentin Thommen,[1] Julien Hurbain,[2] and Benjamin Pfeuty[2]

[1)]*Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, UMR9020-U1277 - CANTHER - Cancer Heterogeneity Plasticity and Resistance to Therapies, F-59000 Lille, France*

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

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

(Dated: 4 October 2022)

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

Keywords: Metabolic flux analysis, Flux balance analysis, Metabolism, Metabolic network 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- 13C]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 13C-MFA (Niedenführ, Wiechert, and Nöh, 2015; Allen and Young, 2020; Antoniewicz, 2021). Despite its limitations, 13C-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 13C dynamic MFA methods (13C-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 bibi 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 bibi reactions in the nonoxidative PPP and where 13C 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

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

2. RESULTS

2.1. Stochastic Simulation Algorithm (SSA)

The propagation of labeled atoms through a biochemical network is here described by a sampling approach. The representation of the isotopomer distribution of each chemical species in the network is computed using a finite sample size proportional to its concentration. A user defined parameter Ω corresponds to a reference concentration. For example, a value of $\Omega = 1000$ c/μM indicates that a concentration of 1 μM is represented by 1000 copies of the chemical species, each copy corresponding to a different isotopomer.

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

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

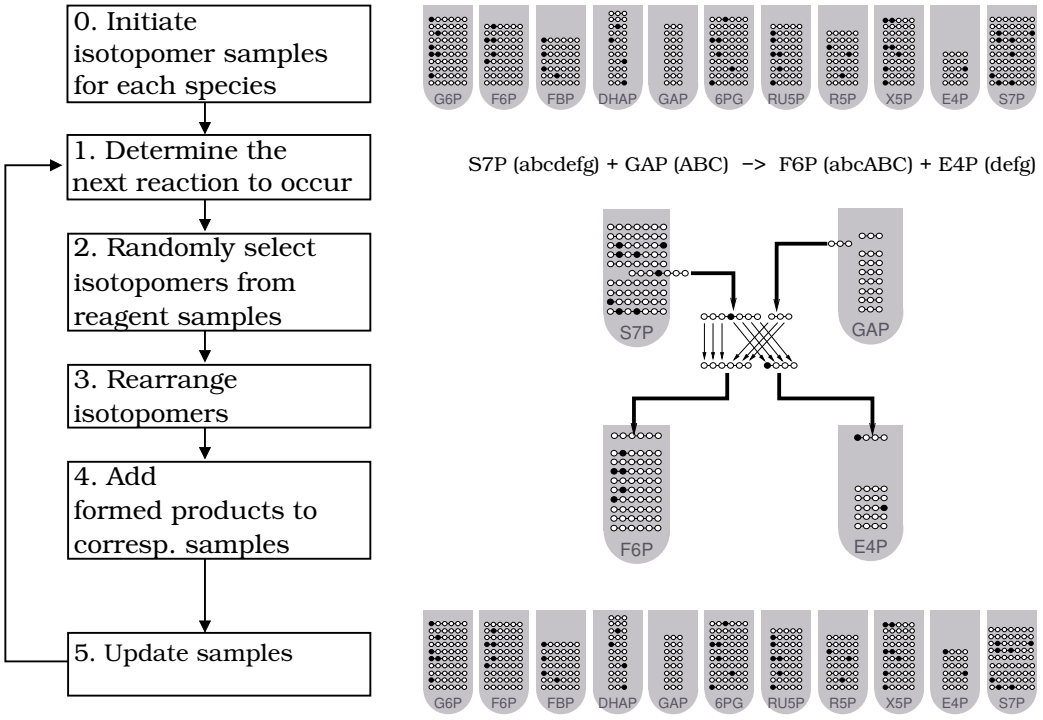

FIG. 1. General principle of the stochastic simulation algorithm (left part) and example of one iteration realization (right part). For each chemical species, the diversity of isotopomers is represented by a sample. The time evolution consists in determining the next reaction from the flux values, and 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.

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

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

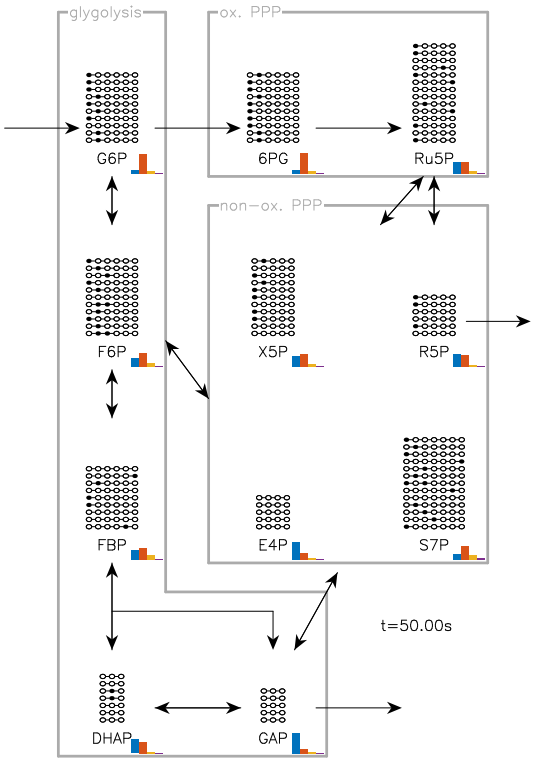

FIG. 2. Atom rearrangements in metabolic reactions. The upper glycolytic pathway supplemented by the pentose phosphate pathway provides an example of isotope labeling network. For each chemical species (G6P, F6P, FBP, DHAP, GAP 6PG, Ru5P, X5P, R5P, E4P, S7P) subsamples of the isotopomer samples are displayed 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 isopotomer 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.

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

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

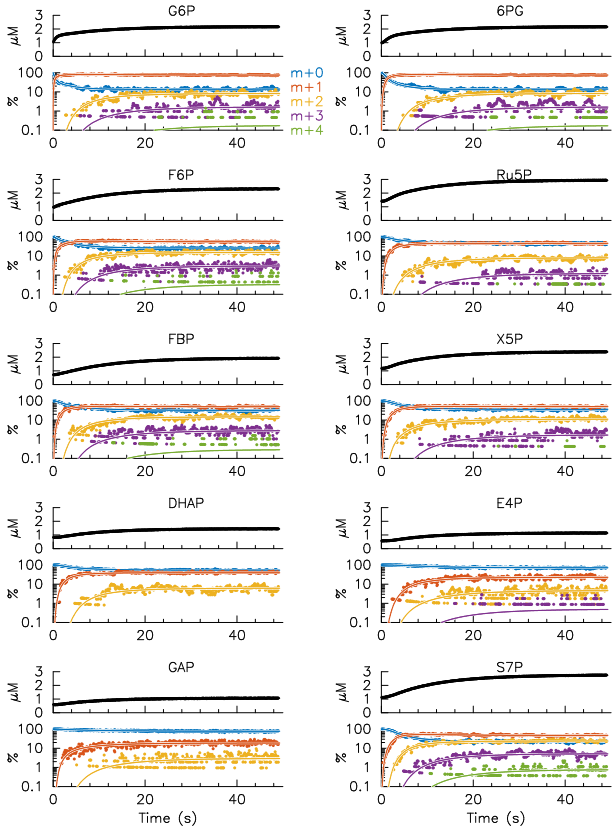

FIG. 3. Concentration and mass isotopomer dynamics in nonstationary conditions. Stochastic trajectories computed with SSA (dots) are compared with deterministic trajectories computed with DSA for control (solid lines), and corresponds to calculation presented in Figure 1 (same network, same condition). 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).

temporal smoothing of the stochastic evolution.

2.2. SSA Computational Performance

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

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

2.3. SSA in 13C-DMFA

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

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

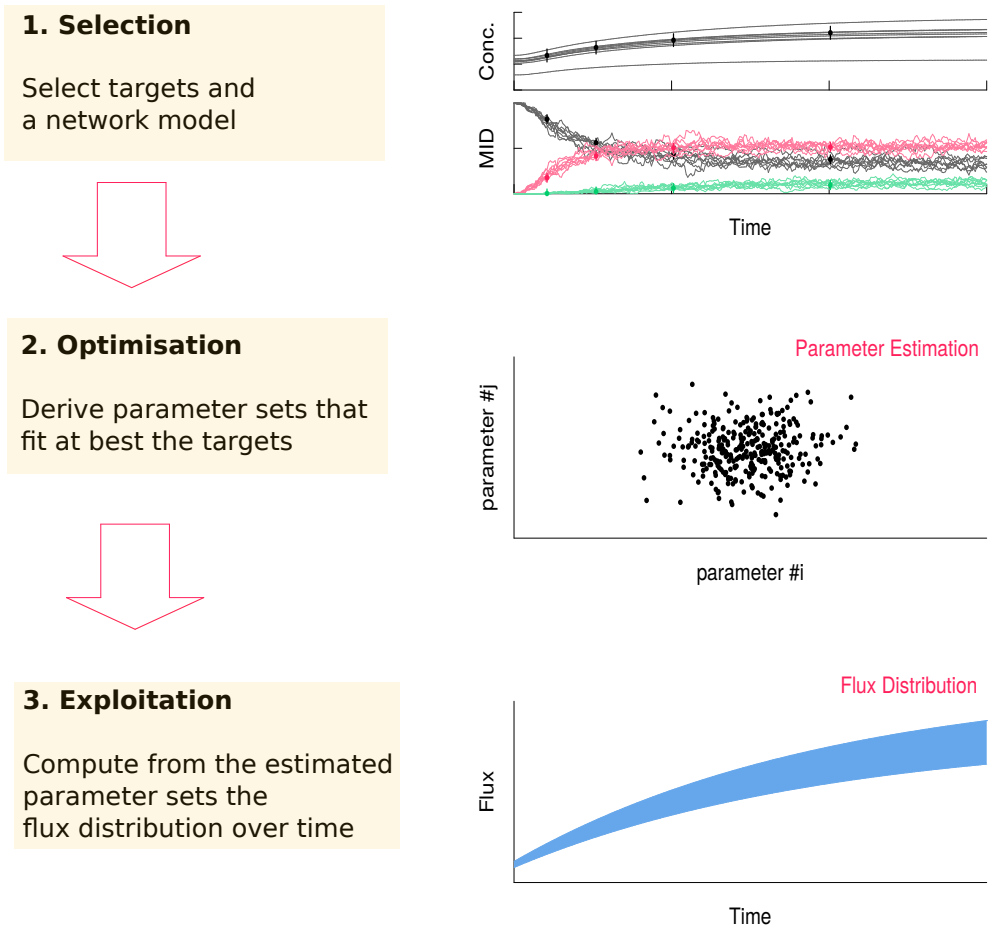

FIG. 4. 13C-DMFA generic workflow using SSA. Metabolite concentration as well as mass isotopomer distribution (MID) are the targets of parameter estimation procedure. Parameter sensitivity analysis provides a list of points in the parameter space that accurately describes targets. The range of flux dynamics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

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

2.4. SSA in 13C-NMFA or 13C-MFA

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

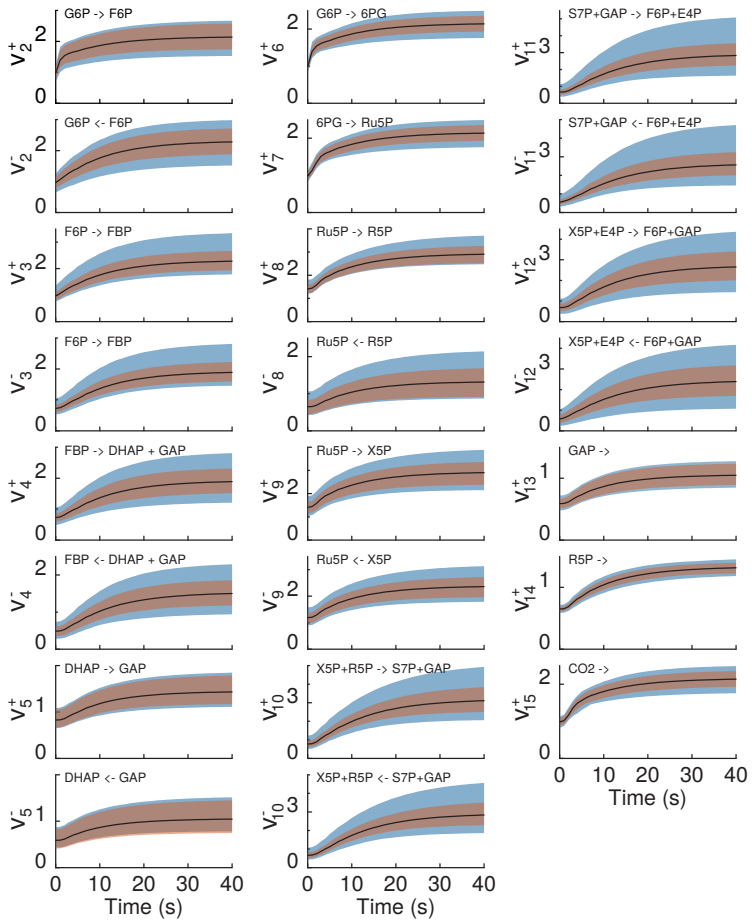

FIG. 5. Directional flux areas compatible with mass isotopomer targeting. The areas are filled with a color code. Blue areas correspond fit strategy 1; red, to fit strategy 2. Solid black lines correspond to the exact solution. $v_i^{+/-}$ correspond to the flux of reaction number i in Table 1 in the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

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

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

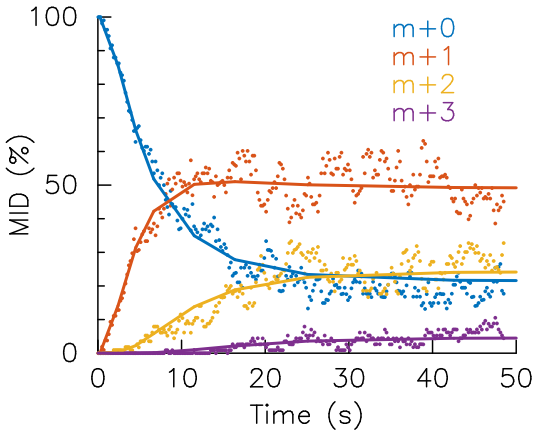

FIG. 6. Comparison of the MIDs dynamics for the S7P variable calculated with the SSA (dots) and INCA software (solid lines). The simulated network corresponds to the reactions listed in Table 1 with equilibrium fluxes (13C-NMFA). The color code is indicated on the figure.

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

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

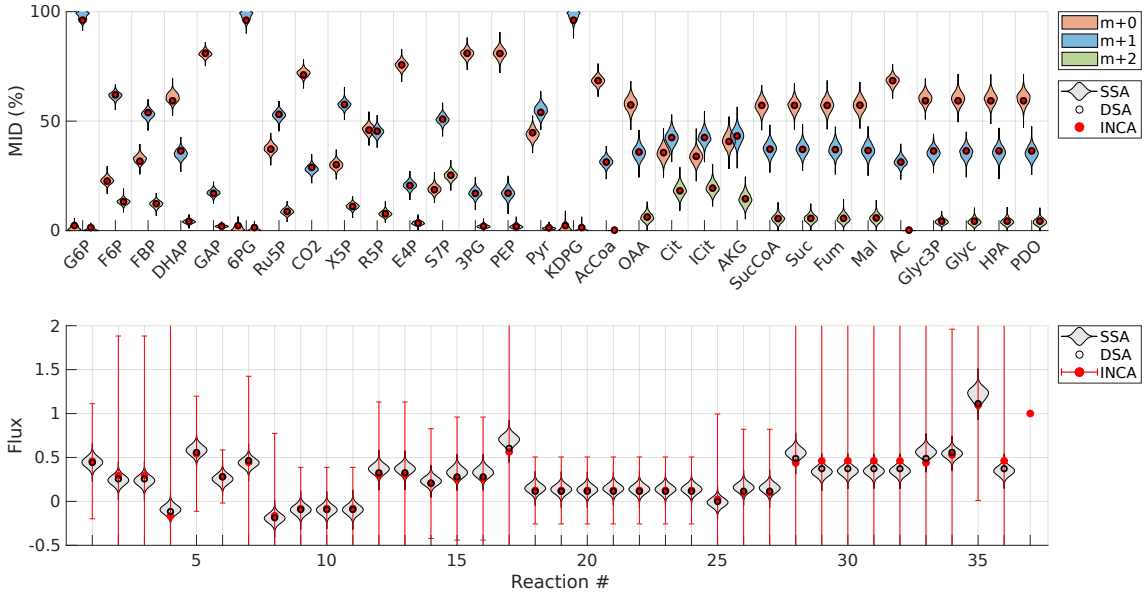

FIG. 7. 13C-MFA with SSA or INCA. A 13C-MFA test performed with SSA coupled to an MCMC method and with INCA on generated MID targets and with a chosen absolute error of 2.5%. Distribution of target fits for each variable (top panel) and estimation of net fluxes (bottom panel). The color code is indicated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

3. THEORY

3.1. Chemical master equation model for isotopic labeling networks

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

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

3.1.1. Isotopomer index and addressing operators

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

If the permutation rule is known, one can define addressing operators that compute the isotopomer 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

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

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

compute the product index from reactant index $n_{\mathrm{S7P}} = (s_{a,1}, \ldots, s_{a,7})_q$ and $n_{\mathrm{GAP}} = (s_{b,1}, s_{b,2}, s_{b,3})_q$. In the same manner,

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

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

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

3.1.2. Chemical master equation description

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

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

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

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

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

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

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 nth-isotopomer of the mth 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,

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

where N denotes the stoichiometry matrix, v_k $k \in [0,K]$ the concentration-dependent flux of reagents, and $\Phi^k_{m,n}(\rho)$ the flux fraction describing the permutation rules of chemical reaction satisfying $\sum_n \Phi^k_{m,n}(\rho) = 1$. Algorithm to simulate Eq. 7 with standard Runge-Kutta-Fehlberg method 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 operations 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, respectively. In that case, the reaction index k is omitted and the flux fraction reads,

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

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

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

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

where L_c and L_d are the number of isotopomers of C and D species. If B is an empty set then $\rho_{b,x} = 1$, and if D is an empty set Eq. (11) is useless. If the reaction is an output reaction, then 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 reactions. As mentioned, the input reactions must be treated separately since the reactants are not variables but parameters. We consider here input reactions of simple form $\emptyset \rightarrow C$ and we note I_n 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

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

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

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

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

3.3. Derived Stochastic Simulation Algorithm (SSA)

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

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

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

Init: Compute the reaction time for all reactions

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

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

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

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

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

and iterate to Step 1

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

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

algorithm to study the stochastic fluctuation of the chemical reactions, it would be enough to use 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 reaction time computations. In the latter case, a realistic estimate for the Ω parameter value must be used.

4. METHODS

4.1. Code availability and computer simulation

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

4.2. Goodness of Fit

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

4.3. Metabolic Network

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

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

TABLE II. List of chemical reactions and reaction speed used to simulate the 13C 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 13C-DMFA of the illustrative case of the pentose phosphate pathways for which 13C-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 13C-MFA (Hurbain *et al.*, 2022), 13C-NMFA or 13C-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 13C 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 13C labeling data become an achievable goal. However, SSA can still be used

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

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

PAS VRAIEMENT A PROPOS, SI ?. Luo, Shen, and Xie (2022)

ACKNOWLEDGEMENTS

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

FUNDING

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

REFERENCES

Allen, D. K. and Young, J. D., "Tracing metabolic flux through time and space with isotope labeling experiments," Current opinion in biotechnology **64**, 92–100 (2020).

Antoniewicz, M. R., "Methods and advances in metabolic flux analysis: a mini-review," Journal of industrial microbiology and biotechnology **42**, 317–325 (2015a).

Antoniewicz, M. R., "Parallel labeling experiments for pathway elucidation and 13c metabolic flux analysis," Current opinion in biotechnology **36**, 91–97 (2015b).

Antoniewicz, M. R., "A guide to metabolic flux analysis in metabolic engineering: methods, tools and applications," Metabolic engineering **63**, 2–12 (2021).

Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G., "Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions," Metabolic engineering **9**, 68–86 (2007).

Antoniewicz, M. R., Kraynie, D. F., Laffend, L. A., González-Lergier, J., Kelleher, J. K., and Stephanopoulos, G., "Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of E. coli producing 1, 3-propanediol," Metabolic engineering **9**, 277–292 (2007).

Baxter, C., Liu, J., Fernie, A., and Sweetlove, L., "Determination of metabolic fluxes in a non-steady-state system," Phytochemistry **68**, 2313–2319 (2007).

Bouzier-Sore, A.-K. and Bolaños, J. P., "Uncertainties in pentose-phosphate pathway flux assessment underestimate its contribution to neuronal glucose consumption: relevance for neurodegeneration and aging," Frontiers in Aging Neuroscience **7**, 89 (2015).

Creek, D. J., Mazet, M., Achcar, F., Anderson, J., Kim, D.-H., Kamour, R., Morand, P., Millerioux, Y., Biran, M., Kerkhoven, E. J., *et al.*, "Probing the metabolic network in bloodstream-form trypanosoma brucei using untargeted metabolomics with stable isotope labelled glucose," PLoS pathogens **11**, e1004689 (2015).

Crown, S. B. and Antoniewicz, M. R., "Parallel labeling experiments and metabolic flux analysis: Past, present and future methodologies," Metabolic engineering **16**, 21–32 (2013).

Diaz-Moralli, S., Aguilar, E., Marin, S., Coy, J. F., Dewerchin, M., Antoniewicz, M. R., Meca-Cortés, O., Notebaert, L., Ghesquière, B., Eelen, G., *et al.*, "A key role for transketolase-like 1 in tumor metabolic reprogramming," Oncotarget **7**, 51875 (2016).

Dong, W., Moon, S. J., Kelleher, J. K., and Stephanopoulos, G., "Dissecting mammalian cell metabolism through 13C-and 2H-isotope tracing: Interpretations at the molecular and systems levels," Industrial & Engineering Chemistry Research **59**, 2593–2610 (2019).

Gibson, M. A. and Bruck, J., "Efficient exact stochastic simulation of chemical systems with many species and many channels," The journal of physical chemistry A **104**, 1876–1889 (2000).

Gillespie, D. T., "Exact stochastic simulation of coupled chemical reactions," The journal of physical chemistry **81**, 2340–2361 (1977).

Gillespie, D. T., "A rigorous derivation of the chemical master equation," Physica A: Statistical Mechanics and its Applications **188**, 404–425 (1992).

Gillespie, D. T., "The chemical langevin equation," The Journal of Chemical Physics **113**, 297–306 (2000).

Gillespie, D. T., "Approximate accelerated stochastic simulation of chemically reacting systems," The Journal of chemical physics **115**, 1716–1733 (2001).

Hartline, C. J., Schmitz, A. C., Han, Y., and Zhang, F., "Dynamic control in metabolic engineering: Theories, tools, and applications," Metabolic engineering **63**, 126–140 (2021).

Heinonen, M., Osmala, M., Mannerström, H., Wallenius, J., Kaski, S., Rousu, J., and Lähdesmäki, H., "Bayesian metabolic flux analysis reveals intracellular flux couplings," Bioinformatics **35**, i548–i557 (2019).

Hurbain, J., Thommen, Q., Anquez, F., and Pfeuty, B., "Quantitative modeling of pentose phosphate pathway response to oxidative stress reveals a cooperative regulatory strategy," bioRxiv

(2022), 10.1101/2022.02.04.478659.

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

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

Lee, M. H., Malloy, C. R., Corbin, I. R., Li, J., and Jin, E. S., "Assessing the pentose phosphate pathway using [2, 3-13c2] glucose," NMR in Biomedicine **32**, e4096 (2019).

Leighty, R. W. and Antoniewicz, M. R., "Dynamic metabolic flux analysis (DMFA): a framework for determining fluxes at metabolic non-steady state," Metabolic engineering **13**, 745–755 (2011).

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

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

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

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

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

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

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

Schmidt, K., Carlsen, M., Nielsen, J., and Villadsen, J., "Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices," Biotechnology and bioengineering **55**, 831–840 (1997).

Schumacher, R. and Wahl, S. A., "Effective estimation of dynamic metabolic fluxes using 13c labeling and piecewise affine approximation: from theory to practical applicability," Metabolites **5**, 697–719 (2015).

Selivanov, V. A., Puigjaner, J., Sillero, A., Centelles, J. J., Ramos-Montoya, A., Lee, P. W.-N., and Cascante, M., "An optimized algorithm for flux estimation from isotopomer distribution in glucose metabolites," Bioinformatics **20**, 3387–3397 (2004).

Stephanopoulos, G., "Metabolic fluxes and metabolic engineering," Metabolic engineering **1**, 1–11 (1999).

Theorell, A., Leweke, S., Wiechert, W., and Nöh, K., "To be certain about the uncertainty: Bayesian statistics for 13c metabolic flux analysis," Biotechnology and bioengineering **114**, 2668–2684 (2017).

Theorell, A. and Nöh, K., "Reversible jump MCMC for multi-model inference in metabolic flux analysis," Bioinformatics **36**, 232–240 (2020).

Thompson, J. M. T. and Stewart, H. B., *Nonlinear dynamics and chaos* (John Wiley & Sons, 2002).

Valderrama-Bahamóndez, G. I. and Fröhlich, H., "Mcmc techniques for parameter estimation of ode based models in systems biology," Frontiers in Applied Mathematics and Statistics **5**, 55 (2019).

Van Kampen, N. G., *Stochastic processes in physics and chemistry*, Vol. 1 (Elsevier, 1992).

Wahl, S. A., Nöh, K., and Wiechert, W., "13 c labeling experiments at metabolic nonstationary conditions: an exploratory study," Bmc Bioinformatics **9**, 1–18 (2008).

Young, J. D., "INCA: a computational platform for isotopically non-stationary metabolic flux analysis," Bioinformatics **30**, 1333–1335 (2014).

Zupke, C. and Stephanopoulos, G., "Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrixes," Biotechnology Progress **10**, 489–498 (1994).