OpenFLUX2: $^{13}$C-MFA modeling software package adjusted for the comprehensive analysis of single and parallel labeling experiments.

Mikhail S Shupletsov$^{1,2,†}$ • Lyubov I Golubeva$^1$ • Svetlana S Rubina$^1$ • Dmitry A Podvyaznikov$^{1,3}$ • Shintaro Iwatani$^{1,††}$ • Sergey V Mashko$^{1,3,4,†}$

$^1$Ajinomoto-Genetika Research Institute, 117545, Moscow, Russian Federation;
$^2$Computational Mathematics and Cybernetics Department, Lomonosov Moscow State University, 119991, Moscow, Russian Federation;
$^3$Department of Theoretical and Experimental Physics, Moscow Physical Engineering Institute (Technical University), 115409, Moscow, Russian Federation;
$^4$Biological Department, Lomonosov Moscow State University, 119991, Moscow, Russian Federation;

‡Correspondence: sergey_mashko@agri.ru

††Present address: Fermentation Group, Process Industrialization Section, Research Institute for Bioscience Products & Fine Chemicals, Ajinomoto Co., Inc., 840-2193, SAGA, Saga-shi, Morodomi-cho, 450 Morodomitsu, Japan
Basic symbols:

\(a, b, c\) – scalar values
\(u, v, w\) – vectors
\(u_i\), resp. \(\{u_i\}\) or \((u)_i\) – the \(i\)-th element of vector \(u\)
\(A, B, C\) – matrices
\(a_{ij}\), resp. \(\{A\}_{ij}\) or \((A)_{ij}\) – the \(i, j\)-th element of matrix \(A\)
\(\{A\}_i\) – the \(i\)-th row of matrix \(A\)
\(\{A\}_j\) – the \(j\)-th column of matrix \(A\)
\(u^T, A^T\) – the transpositions of \(u\) and \(A\)
\(\text{dim}(u), \text{dim}(A)\) – the dimension of \(u\) and \(A\)
\(\text{rank}(A)\) – the rank of \(A\)
\(\text{Null}(A)\) – the null space of \(A\)
\(E[x], E[x]\) - the expected value of random variable \(x\) and random vector \(x\), respectively
\(\Sigma_x\) - the covariance matrix of \(x\)
\(N(\mu, \Sigma)\) - multivariate normal distribution with mean vector \(\mu\) and covariance matrix \(\Sigma\)
\(N(0, 1)\) - standard normal distribution
\(J_f(\hat{\theta})\) - the Jacobian matrix that equals \(\partial f(\theta)/\partial \theta\) evaluated at \(\theta = \hat{\theta}\)
\(H_{SSR}(\hat{\theta})\) - the Hessian matrix of the scalar function \(SSR\) with respect to variables \(\theta\), evaluated at \(\theta = \hat{\theta}\)
\(\chi^2\) – chi-squared distribution
\(\chi^2_{p} \mid \alpha\) – the value of the inverse of the cumulative distribution function of the \(\chi^2\) with \(p\) degrees of freedom at a significance level of \(\alpha\)
\(\chi^2_{\gamma} (p)\) – the same value as previous, but at a confidence level of \(\gamma\), where \(\gamma = 1 - \alpha\).
CONTENT

SF-1.1. Stoichiometric model and assignment of free fluxes 4
SF-1.2. Flux inequality constraints 8
SF-1.3. Flux calculations using regression analysis 11
SF-1.4. Statistical tests of model adequacy 17
SF-1.5. Local linearized statistical approximations 20
SF-1.6. Nonlinear search of flux confidence intervals 25
SF-1.7. Determination of flux confidence intervals using a Monte-Carlo-based approach 31

SF-1 Appendix 1. The special form of the null-space matrix in OpenFLUX(2) 40
SF-1 Appendix 2. Newton’s method of the constrained NLLSP solution 45
SF-1 Appendix 3. Summing elements in an arbitrary row of the contribution matrix 48

References cited in Supplementary File 1 49
SF-1.1. Stoichiometric model and assignment of free fluxes

In $^{13}$C-MFA, a stoichiometric model is based on a mass balance equation for metabolic quasi-steady state conditions:

$$ S \cdot v = 0 , \quad (S-1.1.1) $$

where $S$ is a $(m \times n)$-dimensional, i.e., $\text{dim}(S) = (m \times n)$, stoichiometric matrix with $s_{i,j}$ coefficient of the $i$-th, $i=1,2,\ldots, m$, balanceable metabolite in the $j$-th, $j=1,2,\ldots, n$, reaction, and $v=(v_1, v_2, \ldots, v_n)^T$ is a flux column-vector, $\text{dim}(v) = (n \times 1)$. The stoichiometric model is automatically generated by OpenFLUX(2) from the user-supplied reaction data.

Typically, realistic models are underdetermined because the number of fluxes normally exceeds the number of metabolites ($m < n$) [S-1, S-2]. Under this assumption, the general solution to Eq. $(S-1.1.1)$ can be found e.g., using the method of Gauss-Jordan elimination and partial pivoting that result in the conversion of the $S$ matrix with $r = \text{rank}(S) \leq m$ to the following form:

$$ S' = \begin{pmatrix} I & M \\ 0 & 0 \end{pmatrix} , \quad (S-1.1.2) $$

where $I$ denotes the identity matrix, $\text{dim}(I) = (r \times r)$, and $M$ is the matrix $\text{dim}(M) = (r \times p)$ with $p = (n - r)$, which is obtained from the corresponding part of $S$ after elementary row operations and after column rearrangements with the subsequent reordering of the variables in $v$. Finally, the $u$ flux column-vector, $\text{dim}(u) = (n \times 1)$, is obtained:

$$ u = \begin{pmatrix} V_{dep} \\ 0 \end{pmatrix} , \quad (S-1.1.3) $$
where the upper elements $v_{dep}$, $\dim(v_{dep})=(r \times 1)$, and the lower elements $\theta$, $\dim(\theta)=(p \times 1)$, are assigned as “dependent” and “free” fluxes, respectively. The following equation is derived from Eqs. (S–1.1.1) – (S–1.1.3):

$$S' \cdot u = \begin{pmatrix} I & M \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} v_{dep} \\ 0 \end{pmatrix} = 0. \quad (S–1.1.4)$$

From Eq. (S–1.1.4), according to their assignment, the “dependent” fluxes could be expressed as a linear combination of the “free” fluxes as follows:

$$v_{dep} = -M \cdot \theta. \quad (S–1.1.5)$$

Assigning $N = \begin{pmatrix} -M \\ I \end{pmatrix}$, where $\dim(I) = (p \times p)$ and $\dim(N) = (n \times p)$, the equation for the flux calculation is finally obtained as follows:

$$u = \begin{pmatrix} v_{dep} \\ \theta \end{pmatrix} = \begin{pmatrix} -M \\ I \end{pmatrix} \cdot \theta = N \cdot \theta. \quad (S–1.1.6)$$

Thus, the general solution to the stationary Eq. (S–1.1.4) can be expressed as a linear combination of $p = (n-r)$ independent $\theta$ fluxes (Eq. (S–1.1.6)), where $N$ is the null-space matrix of $S$, $N \equiv \text{Null}(S)$, [S-3, S-4]. In general, the forms of $N$ and the corresponding sets of $\theta$ are not unique for any given $S$, and different methods can be applied for their calculation [S-5]. Nevertheless, $\dim(N)$ and $\dim(\theta)$ are fully determined by the dimension and the rank of the $S$ matrix.

The OpenFLUX(2) software essentially follows the above-described procedure of $N$ and $\theta$ calculation, with the following differences, which improve computational stability: First, matrix $N'$ of the orthonormal basis vectors of the null space of $S$ is obtained using singular value decomposition. Because matrix $N$ is composed of vectors of a specific basis of the null space of $S$, this matrix can be acquired from the $N'$ matrix using the following linear
transforms. First, the $N^*$ matrix is obtained by swapping rows in $N'$, according to the reordering of the initial flux column-vector. Finally, $N$ is acquired from the row-reduced echelon form of $N^*$ using elementary row and column operations.

In OpenFLUX(2), the $v$ flux column-vector is decomposed into the following four different reaction types: bi-directional forward and reverse reactions ($v^+$ and $v^-$) and irreversible reactions that are either free fluxes ($\theta^\text{irr}$) or dependent fluxes ($v^\text{irr}_{\text{dep}}$) [S-6]. Usually, the variable measured effluxes, $v_{\text{eff}}^\text{mea}$, which are constrained within obtained error ranges, are automatically considered to belong to the set of $\theta^\text{irr}$ [S-6, S-7], i.e., $v_{\text{eff}}^\text{mea} = \theta_{\text{eff}}^\text{irr} \subset \theta^\text{irr} \subset \theta$.

According to the user’s choice, the precisely determined effluxes could be considered to be invariant fluxes of the fixed values, i.e., constant parameters. Nevertheless, only the variable effluxes with the experimentally measured standard deviations would be considered in the present study to standardize all provided further computations with the free fluxes, their variances, elements of Jacobian and Hessian matrixes, etc.

Furthermore, some of the rest of the model-based fluxes can also be automatically included in the set of free fluxes. Indeed, let $v_i$ and $v_j$, $1 \leq i \neq j \leq n$, be the elements of column-vector $v$, which correspond to the forward and reverse fluxes, respectively, of one of the $t$ bi-directional reactions. Then, the following equality of the stoichiometric coefficients must be accomplished for the arbitrary $l$-th metabolite, $1 \leq l \leq m$:

$$s_{il} = -s_{lj}. \quad (S-1.1.7)$$

Thus, the $i$-th and $j$-th columns of the $S$ matrix are linearly dependent, indicating that all $t$ columns corresponding to $v^-$, for example, could be
rearranged to the right part of the $S$ matrix at the stage of conversion to the form presented in Eq. ($S$–1.1.2), with the coupled reordering of the position of $v^-$ to the bottom part of the nascent $u$ column-vector in Eq. ($S$–1.1.3). Thus, all $t$ reverse fluxes would be assigned as free fluxes. The residual $(p-t)$ essential free fluxes must be chosen after finalizing the generation of $S'$ in Eq. ($S$–1.1.2). If all $v^-$ fluxes generate a linearly independent set, then Eq. ($S$–1.1.6) and the expressions for $u$, $N$, and $\theta$ can be presented in the following form:

$$
u = \begin{pmatrix} v^- \\ v_{dep}^- \\ 0^- \\ v^+ \end{pmatrix} = \begin{pmatrix} M_{net}^{rev} & 1 \\ M_{net}^{irr} & 0 \\ I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0^- \\ v^- \end{pmatrix} = N \cdot \begin{pmatrix} \theta^{irr} \\ v^- \end{pmatrix} = N \cdot \theta,$$

($S$–1.1.8)

corresponding calculations with introducing of the matrixes $M_{net}^{rev}$ and $M_{net}^{irr}$ definitions are presented in SF-1. Appendix 1.

All free fluxes must be chosen by the user or, alternatively, could be automatically assigned by OpenFLUX(2) (with the default inclusion of $v^-$ and $\theta_{eff}^{irr}$ as a subset of $\theta$). In the former case, the software toolboxes diagnose the user’s choice of the $\theta$ candidates and edit this set, if necessary, due to the computer-based determination of matrix $N$ and to the corresponding set of free fluxes. The selected $\theta$ candidates are verified for their ability to generate a feasible set of non-negative $u$ fluxes according to Eq. ($S$–1.1.6), which are subjected to the flux inequality constraints (see SF-1.2).

**SF-1.2. Flux inequality constraints**
Because every bi-directional reaction is presented by a pair of forward and reverse fluxes in OpenFLUX(2), the following obvious constraints must be applied to the calculation of fluxes:

\[ \mathbf{u}(\theta) = \left( \begin{array}{c} V_{dep} \\ \theta \\ \nu \end{array} \right) \geq 0. \]  

In OpenFLUX2, \( \theta = \left( \begin{array}{c} \theta^{irr} \\ v^{+} \end{array} \right) \) values are non-negative and obey additional constraints. First, \( \theta^{irr} \) is upward bounded by the net flux in the system. Next, it is supposed [S-3, S-8] that a range is given for the variable \( \theta^{irr} \), which is measured with the normally distributed experimental error, and that this range depends on the following efflux confidence intervals:

\[ l_{b_{eff}}^{irr} = \max \left\{ 0, q_{meca}^{eff} - \sigma_{meca}^{eff} \cdot \sqrt{\chi^2_\gamma(1)} \right\} \leq \theta_{off}^{irr} \leq q_{meca}^{eff} + \sigma_{meca}^{eff} \cdot \sqrt{\chi^2_\gamma(1)} = u_{b_{eff}}^{irr}, \]  

where \( q_{meca}^{eff} \) and \( \sigma_{meca}^{eff} \) are the experimentally obtained values of \( \theta^{irr} \) and their measurement variances, respectively; and \( \chi^2_\gamma(1) \) is the value of the inverse of the cumulative distribution function of the \( \chi^2 \)-distribution, with one degree of freedom at a confidence level of \( \gamma \) (i.e., \( \gamma = 0.95 \) for the 95% confidence level).

Thus, \( 100 \times \gamma \% \) from all possible experimental observations of \( \theta_{off}^{irr} \) would be dispersed in the range defined by Ineq. (S-1.2.2).

For the residual (e.g., intracellular) non-negative irreversible free fluxes, \( \theta_{in}^{irr} \), the upper bound, \( u_{b_{in}}^{irr} \), is, by default, 15 times higher than the maximum substrate uptake rate, \( 15 \times \max q_{meca}^{eff} \) [S-6]. In OpenFLUX2, it is possible to augment the range of the \( \theta_{off}^{irr} \) flux variation constrained by Ineq. (S-1.2.2) up to the bounds typical for the \( \theta_{in}^{irr} \) fluxes. This option could be preferable for flux
calculations when the \( q_{\text{mea}}^{\text{eff}} \) and/or \( \sigma_{\text{mea}}^{\text{eff}} \) experimental parameters could be inadequately estimated.

In turn, the \( \mathbf{v}^+ \) fluxes are non-negative but have no natural upper limit. Nevertheless, in OpenFLUX2, the user can assign an artificial upper limit for \( \mathbf{v}^+ \) (e.g., \( \max \mathbf{v}^+ = 200 \times \max q_{\text{mea}}^{\text{eff}} \)). Moreover, a compactification operation is always used by OpenFLUX(2) to transform a physical \( 0 \leq \mathbf{v}^+ < \infty \) scale to a numerical \( 0 \leq \mathbf{v}^+^{(0,1)} < 1 \) scale [S-9] to improve the calculation efficiency [S-3, S-8]. In this situation, \( \mathbf{v}^{\cdot(0,1)} = \frac{\mathbf{v}^+}{\mathbf{P} + \mathbf{v}^+} \) – is a conversion of \( \mathbf{v}^+ \) to the new \( \mathbf{v}^{\cdot(0,1)} \) vector, where \( \mathbf{P} \) denotes a vector of scaling factors with values that could be proportional to \( \max q_{\text{mea}}^{\text{eff}} \) [S-6, S-10]. Thus, in OpenFLUX2, the \( \mathbf{\theta} \) flux-vector and its compactification analog, \( \mathbf{\theta}^{\text{comp}} = \left( \begin{array}{c} \theta_{\text{irr}}^{\text{eff}} \\ \theta_{\text{m}}^{\text{rr}} \end{array} \right) \), is constrained by lower and upper bounds. Indeed, the corresponding constraints for \( \mathbf{\theta} \), in particular, could be the following:

\[
0 \leq \mathbf{b}_\theta = \begin{pmatrix} \mathbf{b}_{\text{irr}}^{\text{eff}} \\ 0 \\ 0 \end{pmatrix} \leq \mathbf{0} = \begin{pmatrix} \mathbf{0}_{\text{irr}}^{\text{eff}} \\ \mathbf{0}_{\text{m}}^{\text{rr}} \end{pmatrix} \leq \mathbf{u}_\theta = \begin{pmatrix} \mathbf{u}_{\text{irr}}^{\text{eff}} \\ \mathbf{u}_{\text{m}}^{\text{rr}} \end{pmatrix} = \mathbf{v}^+, \quad \mathbf{\theta} \in \mathbf{R}^p \equiv [\mathbf{b}_\theta, \mathbf{u}_\theta],
\]

where \( \mathbf{R}^p \) indicates a feasible domain in \( p = (n-r) \)-dimensional space of the \( \mathbf{\theta} \) free flux variation.

Simultaneously, the inequality \( \mathbf{v}_{\text{dep}}(\mathbf{\theta}) = -\mathbf{M} \cdot \mathbf{\theta} \geq 0 \) can not always be satisfied for all \( \mathbf{\theta} \in \mathbf{R}^p \) [S-3]. Thus, the narrowed domain, \( \mathcal{R}^p \equiv [\mathbf{b}_\theta^*, \mathbf{u}_\theta^*] \subseteq \mathbf{R}^p \), could formally be defined as a set of all \( \mathbf{\theta} \in \mathbf{R}^p \) for which only the non-negative values of the \( \mathbf{v}_{\text{dep}}(\mathbf{\theta}) \) fluxes are generated.
Let \( \mathbf{e}_k = (e_{1,k}, e_{2,k}, \ldots, e_{r,k})^T \) and \( k = (1, 2, \ldots, r) \) be a set of column-vectors, \( \text{dim}(\mathbf{e}_k) = (r \times 1) \), where the \( e_{i,k} \) entry of each \( \mathbf{e}_k \) is the following:

\[
e_{i,k} = \begin{cases} 
1, & \text{if } i = k \\
0, & \text{if } i \neq k 
\end{cases} \quad (S-1.2.4).
\]

Then, two linear programming problems must be solved to provide the dependent flux variability analysis, i.e., to establish the available lower and upper bounds (\( \mathbf{l}_b_{v_{dep}} \) and \( \mathbf{u}_b_{v_{dep}} \), respectively) for non-negative variables of the \( \mathbf{v}_{dep} \) flux:

\[
\{ \mathbf{l}_b_{v_{dep}} \}_k = \min \left\{ \mathbf{e}_k^T \cdot (-\mathbf{M}) \cdot \theta \mid \theta \in \mathbb{R}^p \subseteq \mathbb{R}^p \right\} \geq 0,
\]

\[
\{ \mathbf{u}_b_{v_{dep}} \}_k = \max \left\{ \mathbf{e}_k^T \cdot (-\mathbf{M}) \cdot \theta \mid \theta \in \mathbb{R}^p \subseteq \mathbb{R}^p \right\} = \min \left\{ \mathbf{e}_k^T \cdot \mathbf{M} \cdot \theta \mid \theta \in \mathbb{R}^p \subseteq \mathbb{R}^p \right\} \geq 0. \quad (S-1.2.5)
\]

Therefore, only \( \theta \) values that belong to the non-empty \( \mathbb{R}^p \) domain are finally considered by OpenFLUX2 (see SF-1.1.). Thus, the range of the feasible variation of all fluxes, which is constrained by stoichiometric and experimental data, is established:

\[
\mathbf{0} \leq \mathbf{l}_b = \begin{pmatrix} \mathbf{l}_b_{v_{dep}} \\ \mathbf{l}_b_{0} \end{pmatrix} \leq \mathbf{u}(\theta) = \begin{pmatrix} \mathbf{v}_{dep} \\ \mathbf{0} \end{pmatrix} \leq \begin{pmatrix} \mathbf{u}_b_{v_{dep}} \\ \mathbf{u}_b_{0} \end{pmatrix} = \mathbf{u}_b, \text{ i.e., } \mathbf{u}(\theta) \in \mathbb{R}^n \equiv [\mathbf{l}_b, \mathbf{u}_b]. \quad (S-1.2.6)
\]

These constraints formally determine the feasible domain \( \mathbb{R}^n \) in \( n \)-dimensional space for all flux values estimated using the iterative fitting procedure (SF-1.3.) and for potential bounds of the flux confidence intervals (SF-1.5. – SF-1.7.). \( \mathbb{R}^p \) indicates a feasible domain in \( p = (n-r) \)-dimensional space of the \( \theta \) free flux variation. The direct establishment of the \( \mathbb{R}^n \) parameters is necessary for providing, probably, an optimal iterative fitting procedure with the consequently chosen uniformly distributed points \( \mathbf{u}_i(\theta) \), \( i = 1, 2, \ldots, K \) for each from \( K \geq 100 \) trials. At present, the algorithm generating the uniformly
distributed points has not been implemented in OpenFLUX2 yet (but already realized in high-performance 13CFLUX2 software [S-11]). In OpenFLUX2, the random points \( \theta_i, i=1,2,...,K \) from the feasible \( \mathbb{R}^p \)-domain are chosen for the consequent trial, but only those \( \theta \), are used in the fitting procedure that satisfy the \((S-1.2.1)\) inequality.

**SF-1.3. Flux calculations using regression analysis**

Let \( x^{\text{mea}} \equiv \left( \begin{array}{c} x^{\text{mea}}_{\text{MID}} \\ v^{\text{mea}}_{\text{eff}} \end{array} \right) \), \( \dim(x^{\text{mea}}_{\text{MID}}) = (w_{\text{MID}} \times 1) \), \( \dim(v^{\text{mea}}_{\text{eff}}) = (w_{\text{eff}} \times 1) \), \( \dim(x^{\text{mea}}) = (w \times 1) \), where
\[
w = (w_{\text{MID}} + w_{\text{eff}}),
\]
be a variable column-vector of experimentally measured data (containing both labeling measurements, \( x^{\text{mea}}_{\text{MID}} \), and the effluxes \( v^{\text{mea}}_{\text{eff}} \)) with the corresponding measurement noise vector \( \delta^{\text{mea}} = \left( \begin{array}{c} \delta^{\text{mea}}_{\text{MID}} \\ \delta^{\text{mea}}_{\text{eff}} \end{array} \right) \). If \( \langle x \rangle \) is the unknown in reality true value, then:
\[
\langle x \rangle = x^{\text{mea}} + \delta^{\text{mea}} \quad (S-1.3.1).
\]

Usually (see e.g., [S-11]), it is assumed that the noise vector, \( \delta^{\text{mea}} \), is normally distributed with an expectation vector \( E[\delta^{\text{mea}}] = 0 \), where \( E \) is the expected value operator, with a measurement variance-covariance matrix, \( \Sigma_{\times} \), which is expressed as follows:
\[
\Sigma_{\times} = E[(x^{\text{mea}} - \langle x \rangle) \cdot (x^{\text{mea}} - \langle x \rangle)^T] = E[\delta^{\text{mea}} \cdot (\delta^{\text{mea}})^T],
\quad (S-1.3.2)
\]
i.e., \( \delta^{\text{mea}} \sim \mathcal{N}(0, \Sigma_{\times}) \).

Notably, the correct bias estimation and the removal of the background signals from “raw” MS-mediated data, in particular, are crucial for the
evaluation of \( \mathbf{x}_{\text{MID}}^{\text{mea}} \), which is consistent with the assumed normality of the \( \hat{\mathbf{\delta}}_{\text{MID}}^{\text{mea}} \) distribution [S-12].

OpenFLUX(2) assumes individual measurements to be pair-wise, independent, and uncorrelated. Under these assumptions, \( \Sigma_x \) can be considered diagonal, with squared measurement-variances as diagonal elements as follows:

\[
\left( \Sigma_x \right)_{ij} = \mathbb{E}[(\delta_i^{\text{mea}} \cdot (\delta_j^{\text{mea}})^\top)] = \left( \sigma_i^{\text{mea}} \right)^2 \\
\Sigma_x = \Sigma_x \left( \sigma_i^{\text{mea}} \right) = \mathbf{I} \cdot \left( \sigma_i^{\text{mea}} \right)^2 
\]

where \( \mathbf{I} \) is the identity matrix \( \dim(\mathbf{I}) = (w \times w) \), and \( \left( \sigma_i^{\text{mea}} \right)^2 = \left( \begin{pmatrix} \sigma_{i,\text{MID}}^{\text{mea}} \\ \sigma_{i,\text{MID}}^{\text{mea}} \end{pmatrix} \right)^2 \) is the column-vector of the squared measurement variances \( \dim \left( \left( \sigma_i^{\text{mea}} \right)^2 \right) = (w \times 1) \).

Let \( \mathbf{x}_{\text{input}} \) be a column-vector of the labeling data of input substrate(s) and \( f_i(\mathbf{x}_{\text{input}}, \mathbf{\theta}), i = 1, 2, \ldots, w_{\text{MID}} \), is parameterized by the \( \mathbf{x}_{\text{input}} \) function, which describes the relation between the free fluxes \( \mathbf{\theta} \) and the \( i \)-th MID fraction of the intracellular metabolite (e.g., proteinogenic amino acids). Then, the following isotopomer balance model could be generated as a nonlinear \( f \) vector-function, which maps \( \mathbf{x}_{\text{input}} \) and \( \mathbf{\theta} \) to a column-vector of simulated measurements data \( \mathbf{x}_{\text{calc}} \), \( \dim(\mathbf{x}_{\text{calc}}) = (w \times 1) \):

\[
\mathbf{x}_{\text{calc}} = f(\mathbf{x}_{\text{input}}, \mathbf{\theta}) = \left( f_1(\mathbf{x}_{\text{input}}, \mathbf{\theta}), \ldots, f_{w_{\text{MID}}}(\mathbf{x}_{\text{input}}, \mathbf{\theta}), \left[ \mathbf{\theta}_{\text{eff}} \right]_1, \ldots, \left[ \mathbf{\theta}_{\text{eff}} \right]_{w_{\text{ef}}} \right)^\top. 
\]

The resulting \( \mathbf{x}_{\text{calc}} \) vector corresponds to the \( \mathbf{x}_{\text{mea}} \) measurements vector. In OpenFLUX(2), \( f_i(\mathbf{x}_{\text{input}}, \mathbf{\theta}), i = 1, 2, \ldots, w_{\text{MID}} \), is the value of an EMU-based function that describes the relation between fluxes and the \( i \)-th MID fraction of the intracellular metabolite, which is calculated using the decomposition algorithm,
with the sequential identification of all reactant EMUs that contribute to the $i$-th EMU product [S-13]. The nonlinearity of the generated $f_i(x^{\text{input}}, \theta)$ dependent variables with respect to the flux is based specifically on the fact that the full set of equations contains isotopomer abundance obtained due to unimolecular (linear) and condensation (bilinear) EMU reactions in the system [S-4, S-14]. Therefore, the nonlinear flux-dependent coefficients for the expression of the targeted EMU MIDs contain convolution terms involving lower level EMU variables. The automatic generation of the isotopomer balance model according to the EMU decomposition framework [S-13] using the algorithms applied in OpenFLUX(2) has been described previously [S-6].

Based on the generated isotopomer balance model $f$, the following $\varepsilon$ vector function is parameterized by $x^{\text{input}}$ and $x^{\text{mea}}$, which map values of $\theta$ to the vector of residual errors between calculated values and experimental data, respectively:

$$
\varepsilon(x^{\text{input}}, x^{\text{mea}}, \theta) = x^{\text{calc}}(x^{\text{input}}, \theta) - x^{\text{mea}} = f(x^{\text{input}}, \theta) - x^{\text{mea}}.
$$

(5.3.1)

Then, the measurement-variance-weighted sum of squared residual errors is assigned as the following objective scalar-function of the $\theta$ variables and is parameterized by $x^{\text{input}}$, $x^{\text{mea}}$, and $\sigma^{\text{mea}}$:

$$
SSR_{x^{\text{SLE}}}^{\text{SLE}}(x^{\text{input}}, x^{\text{mea}}, \sigma^{\text{mea}}, \theta) = \frac{1}{2} \cdot (\varepsilon(x^{\text{input}}, x^{\text{mea}}, \theta))^T \cdot \Sigma_{\text{x}}^{-1} \cdot (\sigma^{\text{mea}}) \cdot \varepsilon(x^{\text{input}}, x^{\text{mea}}, \theta),
$$

(6.3.1)

where $\Sigma_{\text{x}}^{-1}$ is the inverse of the $\Sigma_{\text{x}}$ matrix. In the case of the diagonal form of $\Sigma_{\text{x}}$ (see Eq. (5.1.3)), the applied weighting procedure normalizes the squared residual errors between the observed and simulated measurements using the squared measurement variances $(\sigma^{\text{mea}})^2$. The coefficient "$1/2$" is usually introduced in the definition of the $SSR_{x^{\text{SLE}}}$-like objective functions [S-3, S-8, S-
to compensate for the coefficient “2”, which appears when the squared objective function is differentiated (see **SF-1 Appendix 2**). In contrast, to provide statistical tests for model adequacy, the \( \text{N}(0,1) \) normal distribution of the individual residuals and the correspondence of their sum to \( \chi^2 \)-distribution at the point of convergence would be necessary to verify the \( \text{SSR}_{\text{SLE}} \)-like objective function without \( \frac{1}{2} \) as the coefficient. Thus, the \( \Xi = 2 \cdot \text{SSR}_{\text{SLE}} \) function would be particularly assigned for these purposes in **SF-1.4**. Namely, this \( \Xi(x_{\text{input}}, x_{\text{mea}}, \sigma_{\text{mea}}, \theta) \) function (instead of \( \text{SSR}_{\text{SLE}} \)) is assigned as an objective function in several publications (e.g., [S-6, S-16, S-17, S-18 or S-4] where the authors do not mention the Hessian matrix of the objective function or omit some details that are essential for its approximation; see **SF-1 Appendix 2**).

Let \( m_{\text{input}} \) be a vector of \( x_{\text{input}} \) values, which has been used in the actual provided SLE, and \( m_{\text{mea}} = \left( m_{\text{mea MID}}, q_{\text{mea eff}} \right) \) is an experimental realization of the \( x_{\text{mea}} \) vector in this SLE. According to Eqs. (S–1.3.3), (S–1.3.5), and (S–1.3.6), the obtained value of the \( \text{SSR}_{\text{SLE}} \) objective function can be presented as follows:

\[
\text{SSR}_{\text{SLE}}(m_{\text{input}}, m_{\text{mea}}, \sigma_{\text{mea}}, \theta) = \frac{1}{2} \left( \sum_{i=1}^{w_{\text{eff}}} \left( \frac{f_i(m_{\text{input}}, \theta) - \left( m_{\text{mea MID}} \right)_i}{\sigma_{\text{mea MID}}[i]} \right)^2 + \sum_{j=1}^{w_{\text{irr}}} \left( \frac{\left( q_{\text{eff}} \right)_j - \left( q_{\text{mea eff}} \right)_j}{\sigma_{\text{eff}}[j]} \right)^2 \right). 
\]  

(S–1.3.7)

The following nonlinear least-squares minimization problem (NLLSP) must be solved to estimate the optimized \( u(\theta) \) parameters, which are subjected to constraints:

\[
\begin{align*}
\min_{\theta \in \mathbb{R}^n} & \text{SSR}_{\text{SLE}}(m_{\text{input}}, m_{\text{mea}}, \sigma_{\text{mea}}, \theta) \\
\text{subject to} & \\
& u(\theta) \in \mathbb{R}^n
\end{align*}
\]  

(S–1.3.8)

An analogous strategy can be used to calculate the optimized \( u(\theta) \) in the PLE, including \( k \) independent LEs. The following \( z \)-like vectors,
\( z = (z_1^T, z_2^T, \ldots, z_k^T)^T \), can be considered for this PLE, where \( z \) denotes the nonlinear vector-function with entries \( (z) \), corresponding to vector-function \( \mathbf{f} \) of \( i \)-th LE \((1 \leq i \leq k)\), with defined vectors \( \mathbf{m}_i^{\text{input}} \), \( \mathbf{m}_i^{\text{mea}} \), and \( \sigma_i^{\text{mea}} \). In the case of the independence of individual LEs in the analyzed PLE, the objective scalar-function of \( \theta \) with \( \mathbf{m}^{\text{input}} = ((\mathbf{m}_1^{\text{input}})^T, (\mathbf{m}_2^{\text{input}})^T, \ldots, (\mathbf{m}_k^{\text{input}})^T)^T \),
\( \mathbf{m}^{\text{mea}} = ((\mathbf{m}_1^{\text{mea}})^T, (\mathbf{m}_2^{\text{mea}})^T, \ldots, (\mathbf{m}_k^{\text{mea}})^T)^T \) and \( \sigma^{\text{mea}} = ((\sigma_1^{\text{mea}})^T, (\sigma_2^{\text{mea}})^T, \ldots, (\sigma_k^{\text{mea}})^T)^T \), which are the vector parameters, could be presented in the following form:

\[
SSR_{z}^{\text{PLE}}(\mathbf{m}_i^{\text{input}}, \mathbf{m}_i^{\text{mea}}, \sigma_i^{\text{mea}}, \theta) = \sum_{i=1}^{k} SSR_{L}^{\text{LE}}(\mathbf{m}_i^{\text{input}}, \mathbf{m}_i^{\text{mea}}, \sigma_i^{\text{mea}}, \theta),
\]

(\( S-1.3.9 \))

and the following constrained NLLSP must be solved to calculate the optimized \( \mathbf{u}(\theta) \) fluxes, which are the variable parameters of the common metabolic model:

\[
\begin{cases}
\min_{\theta \in \mathbb{R}^p} SSR_{z}^{\text{PLE}}(\mathbf{m}_i^{\text{input}}, \mathbf{m}_i^{\text{mea}}, \sigma_i^{\text{mea}}, \theta) \\
\mathbf{u}(\theta) \in \mathbb{R}^n
\end{cases}
\]

(\( S-1.3.10 \))

If individual LEs introduce significant grouping factors or correlation effects in the flux estimation, then hierarchical regression models, i.e., multi-level models, must be used for flux calculations, with the corresponding application of the modified least-squared approaches [S-19]; however, these strategies were not implemented in the presented software.

One of the possible solutions of the NLLSP expressed by Eqs. (\( S-1.3.8 \)) and (\( S-1.3.10 \)) can be obtained using a gradient-based local optimization algorithm [S-20, S-21], i.e., a variation of Newton’s method, which is explained in \textit{SF-1 Appendix 2}.

In OpenFLUX(2), NLLSP Eq. (\( S-1.3.8 \)) for SLEs is solved with the assistance of the MATLAB’s FMINCON function, which is included in the Optimization Toolbox. Analogously, Eq. (\( S-1.3.10 \)) for PLEs is solved by
FMINCON in OpenFLUX2. By default, FMINCON implements a quasi-Newton sequential quadratic programming (SQP) method for constrained, nonlinear optimization using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula, which is based on the linearization of the Jacobian matrix to approximate the Hessian matrix [S-21], instead of the approximation used in SF-1 Appendix 2, Eq. (4–2.9). Notably, from a computational point of view, a quasi-Newton method is preferred to Newton’s approach because the former directly generates an approximation of the inverse Hessian matrix. However, the analytical form of the Hessian matrix is extremely important in the statistical analysis and in the calculation of the contribution matrix and of the D-factor for optimizing the experimental design (see (SF-1.5)).

The applied SQP algorithm of the NLLSP solution can potentially converge to local minima of the objective function [S-20]. Thus, an iterative optimization procedure with a different set of initial \( \theta \in \mathbb{R}^p \) values must be used to search for a global minimum [S-3, S-16]. In the present study, \( K \) reiterative trials, where \( K \geq 100 \), are typically used to obtain the set of optimized flux estimates, \( u_i(\hat{\theta}_i), \quad i = 1, 2, \ldots, K \). Usually, each \( i \)-th optimization trial proceeds through step-wise iterative changes of free flux value, i.e., \( \theta_1, \theta_2, \ldots, \theta_g \), until reaching the minimum of the target objective function. In OpenFLUX(2), by default, the termination criterion is defined by the Karush-Kuhn-Tucker conditions, with a user-provided tolerance of the change in the objective function value (e.g., \( 1 \times 10^{-4} \)) and the constraint violation tolerance (e.g., \( 1 \times 10^{-3} \)). In contrast, termination can occur not only if the minimization criteria are satisfied but also, for example, if the amount of the iterations, \( g \), which are valid during one trial, is exceeded. The result of such a trial cannot be
recognized as an intrinsic solution of the raised minimization problems. In OpenFLUX(2), the reasons for each termination are included in the Optimization report; therefore, only the successfully terminated trials can be considered. The further statistical evaluation of the resulting flux pseudo-populations of the $u_i(\hat{\theta})$, $i=1, 2, ..., K$, can indicate the existence of feasible unique or multiple optimization solution(s).

**SF-1.4. Statistical tests of the model adequacy**

The computation of a set of NLLSP solutions, $u_i(\hat{\theta})$, does not indicate that the user-made model is adequate. A goodness-of-fit analysis of the model is an essential related part of the flux determination using a regression procedure [S-4, S-22].

When the flux model adequately represents measurement data, then the arbitrary $i$-th residuals in the $\Xi(\theta) = 2 \cdot SSR_i(\theta)$ function evaluated at the point of convergence (i.e., $\left( \frac{f(m_{\text{m}}(\hat{\theta}), m_{\text{input}}) - m_{\text{MID}}}{\sigma_{\text{MID}}} \right)_i$, $i=1, 2, ..., w_{\text{MID}}$, and

$$\left( \frac{\hat{\theta}_{\text{irr}} - (q_{\text{irr}})}{\sigma_{\text{irr}}} \right)_j, \ j=1, 2, ..., w_{\text{eff}}$$

are expected to follow $\mathcal{N}(0,1)$ distribution.

Therefore, according to the $\chi^2$-statistics, the null hypothesis, $H_0$, that the $\Xi(\hat{\theta})$ value, which is a stochastic variable, $\xi$, follows a $\chi^2$-distribution at a point of convergence is considered along with the complex alternative, $H_1$, that the specified value has a different distribution. Moreover, the number of degrees of freedom for the $\chi^2$-distribution must be equal to the number of independent measurements $W$ ($W = w$ in the case of SLE, and $W = \sum_{i=1}^{k} w_i$ in the case of PLE,
respectively), minus the number of free fluxes, \( p \), as estimated parameters, i.e., \((W - p)\). Thus, to test the goodness-of-fit, the following statistics are used:

\[
\delta(\xi) = \begin{cases} 
H_0, & \text{if } \chi^2_{a/2}(W - p) < \xi < \chi^2_{1 - a/2}(W - p); \\
H_0, & \text{if } \xi \leq \chi^2_{a/2}(W - p); \text{ however, the model might be overfitted}; \\
H_1, & \text{if } \xi \geq \chi^2_{1 - a/2}(W - p),
\end{cases}
\]

where \( \alpha \) is a significance level used as a certain threshold, e.g., 0.05 for a 95% confidence interval. If the determined \( \xi \) value exceeds the maximum critical value, \( \chi^2_{1 - a/2}(W - p) \), then the \( H_0 \) hypothesis must be rejected because the used flux model does not satisfy the experimental data at the tested significance level. Furthermore, if the acquired value \( \xi \) is lower than the minimum critical value, \( \chi^2_{a/2}(W - p) \), then the weighted difference between the experimentally measured data and their model-based estimations is questionably small or the measurement’s variances have been overestimated. Strictly speaking, the former has not been a cause to reject the hypothesis \( H_0 \) [S-23] but might serve as an indication of model overfitting using the specified set of parameters.

Usually, the effect of overfitting occurs in complex models due to an imbalance between the number of measurements and the number of fitted parameters [S-24].

Initially, to analyze the set of NLLSP solutions obtained in \( K \) trials, \( u_i(\hat{\theta}_i), \ i = 1, 2, ..., K \), it seems logical to confirm the normality of the distribution of the individual residuals, followed by a goodness-of-fit test of the \( \Xi(\hat{\theta}_i) \) values. However, practically, it is convenient to initiate analysis due to the selection of all solutions by the \( \chi^2 \)-mediated criterion. In this case, the normality of the distribution can be analyzed only for the more interesting solutions (e.g., for one \( i \)-th solution corresponding to the minimal value of \( \Xi(\hat{\theta}_i) \) or for all
solutions for which the $H_0$ hypothesis is accepted). This normality test can be extremely important, particularly when no obtained solutions satisfy the criterion of the $H_0$ hypothesis. Indeed, sometimes these residual values reveal systematic deviations, which may show how the experiment or the man-made model can be improved [S-7, S-20]. If several weighted residuals demonstrate a significant lack-of-fit in these statistical tests, then the results of the flux estimation process could be incorrect, most likely due to (i) the corresponding entries of $m^{\text{mea}}$ and/or $\sigma^{\text{mea}}$ being inadequately evaluated and (ii) the use of the wrong flux model structure. In the case of (i), a reasoned modification of the variance value or a deletion of the corresponding residuals from $\Xi(\theta)$ could result in the acquisition of the proper values of $u(\hat{\theta})$ fluxes after the new attempt of the optimization procedure. In the case of (ii), particularly if a significant number of weighted residuals do not satisfy the tested criteria, then the re-evaluation of the flux model is necessary. Notably, the relative importance of residuals that have been chosen as lack-of-fit candidates for discard from $\Xi(\theta)$ can also be quantitatively evaluated from the contribution matrix calculated by OpenFLUX2 (see SF-1.5). If alternative residuals can be chosen for deletion, then primarily the measurements that manifest lower relative importance can be used.

In OpenFLUX2, the proposed $N(0, I)$ distribution of individual measurement variance-weighted residuals in the expression of $\Xi(\theta)$ is analyzed with the assistance of MATLAB’s Statistic Toolbox using the Kolmogorov-Smirnov test. Moreover, a plot of the absolute value of the individual variance-weighted residuals could be drawn. A simple visual inspection of the plot aids in identifying unpredictable high values for the individual residuals/group of
the individual residuals, followed by the re-evaluation of the corresponding measurements with their variances and/or by the re-evaluation of the man-made metabolic model.

**SF-1.5. Local linearized statistical approximations**

Several important statistical properties of the NLLSP solution could be determined using local linearized approximations of the nonlinear $f(x_{\text{input}}, \theta)$ vector-function in the neighborhood of the convergence point \([S-4, S-10, S-12]\). The first of these properties is the flux identifiability analysis, i.e., whether the NLLSP has unique or multiple solutions that can be estimated from the available experimental data.

Assuming that the obtained estimations $\hat{\theta}$ are close to the true free flux values, the nonlinear $f(x_{\text{input}}, \theta)$ vector-function in Eq. \((S-1.3.4)\) can be linearly approximated using the Taylor’s expansion in the neighborhood of $\hat{\theta}$:

$$f(\theta) \approx f(\hat{\theta}) + J_f(\hat{\theta}) \cdot (\theta - \hat{\theta}), \quad \text{\(S-1.5.1\)}$$

where $J_f(\hat{\theta})$ is the Jacobian matrix that equals $\partial f(\theta)/\partial \theta$ evaluated at $\theta = \hat{\theta}$. Let $y = m_{\text{mea}} - [f(\hat{\theta}) - J_f(\hat{\theta}) \cdot \hat{\theta}]$, then the initial nonlinear regression model in Eq. \((S-1.3.5)\) can be approximated using the following linear regression model:

$$y = J_f(\hat{\theta}) \cdot \theta + (-\varepsilon), \quad \text{\(S-1.5.2\)}$$

where $\varepsilon$ (see Eq. \((S-1.3.5)\)) expectedly follows a multivariate normal distribution, i.e., $\varepsilon \sim N(0, \Sigma_{\varepsilon})$, with $\Sigma_{\varepsilon}$ defined in Eq. \((S-1.3.3)\). Regression parameters, $\theta$, of the linear model represented in Eq. \((S-1.5.2)\) can be estimated for the objective function $\Phi(y, \sigma_{\text{mea}}, \theta)$, which is assigned as follows:

$$\Phi(y, \sigma_{\text{mea}}, \theta) = [y - J_f(\hat{\theta}) \cdot \theta]^T \cdot (\Sigma_{\varepsilon})^{-1} \cdot [y - J_f(\hat{\theta}) \cdot \theta], \quad \text{\(S-1.5.3\)}$$
as a solution of the following generalized least squares problem (GLSP):

\[
\theta_{GLSP} = \arg \min_{\theta} \left( [y - J_f(\hat{\theta}) \cdot \theta]^T \cdot (\Sigma_\eta)^{-1} \cdot [y - J_f(\hat{\theta}) \cdot \theta] \right).
\] (S-1.5.4)

The analytical solution of this GLSP can be found as a solution of the following equation:

\[
\frac{\partial \Phi(\theta)}{\partial \theta} = -2 \cdot J_f(\hat{\theta})^T \cdot (\Sigma_\eta)^{-1} \cdot (y - J_f(\hat{\theta})) \cdot \theta = 0.
\] (S-1.5.5)

Equation (S-1.5.5) can be transformed into the following form:

\[
(J_f(\hat{\theta})^T \cdot (\Sigma_\eta)^{-1} \cdot J_f(\hat{\theta})) \cdot \theta = (J_f(\hat{\theta})^T \cdot (\Sigma_\eta)^{-1}) \cdot y,
\] (S-1.5.6)

which results in the linear estimation of the \( \theta \) parameters in the neighborhood of \( \hat{\theta} \) as a particular solution of Eq. (S-1.5.6):

\[
\theta_{GLSP} = ((J_f(\hat{\theta}))^T \cdot (\Sigma_\eta)^{-1} \cdot J_f(\hat{\theta}))^{-1} \cdot ((J_f(\hat{\theta}))^T \cdot (\Sigma_\eta)^{-1} \cdot y,
\] (S-1.5.7),

with the \( \Sigma_\eta \) covariance matrix of the estimated free fluxes as follows:

\[
\Sigma_\eta \equiv \text{cov}(\hat{\theta}) = ((J_f(\hat{\theta}))^T \cdot (\Sigma_\eta)^{-1} \cdot J_f(\hat{\theta}))^{-1}.
\] (S-1.5.8)

Simultaneously, it is known (e.g., see [S-3, S-12]) that a general solution of a linear nonhomogeneous system (such as in Eq. (S-1.5.6)), \( A \cdot x = b \), can be expressed as the sum of a fixed particular solution, e.g., \( (A^T \cdot A)^{-1} \cdot A^T \cdot b \), and a general solution of the homogeneous system, \( A \cdot x = 0 \), i.e., an arbitrary element, \( c \), of the null space, \( c \in \text{Null}(A) \), or \( c = \text{Null}(A) \cdot \beta \), where \( \beta \) is a vector containing arbitrary non-zero entries. Specifically, the solution set to Eq. (S-1.5.6) is:

\[
\theta = \Sigma_\eta \cdot (J_f(\hat{\theta}))^T \cdot (\Sigma_\eta)^{-1} \cdot y + \text{Null}[J_f(\hat{\theta})] \cdot \beta,
\] (S-1.5.9)

where \( \text{Null}[J_f(\hat{\theta})] \) is the null space of the \( J_f(\hat{\theta}) \) matrix. The following can be concluded from the provided free flux linearized approximations (S-1.5.9): these \( \hat{\theta} \) values of the \( \hat{\theta} \) flux vector can be determined uniquely from the
\( \mathbf{y}(\mathbf{x^{meas}}) \) vector, depending on the experimentally measured data for whose estimates the corresponding rows of the \( \text{Null}[\mathbf{J}_i(\hat{\theta})] \) matrix consist of zeros. In contrast, if the entry of \( \text{Null}[\mathbf{J}_i(\hat{\theta})] \) corresponding to \( \hat{\theta}_i \) is non-zero, then infinitively many solutions for \( \hat{\theta}_i \) at the obtained set of experimental data can be expected [S-12]. Thus NLLSP has a unique solution that justifies the available experimental data when the \( \text{Null}[\mathbf{J}_i(\hat{\theta})] \) matrix is empty.

Due to the limited precision of the floating-point numbers, in certain cases numerical estimation of the \( \text{Null}[\mathbf{J}_i(\hat{\theta})] \) can be inaccurate and result in unreliable conclusion of flux identifiability. To evaluate reliability of identifiability analysis in OpenFLUX2, condition number \( \text{cond}(\mathbf{J}_i(\theta)) \) is compared to precision \( \text{eps}(\mathbf{J}_i(\theta)) \) of the floating-point representation of the computed Jacobian values. As a general rule of thumb, if the condition number \( \text{cond}(A) = 10^n \), then up to \( n \) digits of accuracy might be lost in computations. Thus, identifiability results can be considered reliable only when condition number \( \text{cond}(\mathbf{J}_i(\hat{\theta})) \) is within a certain upper bound. Taking into consideration the null space computation method used in MATLAB the following bound was chosen. If \( \text{cond}(\mathbf{J}_i(\hat{\theta})) < \frac{1}{\sqrt{\text{eps}(\mathbf{J}_i(\hat{\theta}))}} \), condition number is considered acceptable and computations are treated as reliable. Otherwise, condition number is too large, Jacobian matrix is considered to be ill-conditioned and results of identifiability analysis are treated as unreliable. Corresponding messages about condition number status are displayed alongside with judgment about identifiability of fluxes, based on computation of \( \text{Null}[\mathbf{J}_i(\hat{\theta})] \).
As observed from the direct comparison of the $\Sigma_\theta$ covariance matrix determination presented in $S-1.5.8$, an approximation of the reverse Hessian matrix, $[H_{SSR}(\theta)]^{-1}$, which is introduced in Appendix 2, Eq. (A2.10), the following equation has to be true:

$$
\Sigma_\theta \approx [H_{SSR}(\theta)]^{-1}.
$$

This relation is usually considered one of the convenient ways to compute the approximate covariance matrix of the estimated parameters [S-23].

From the covariance matrix properties and from the relation between the $\theta$ and $u$ fluxes expressed by Eq. $(S-1.6)$, the covariance matrix for the estimated optimal fluxes, $\Sigma_u$, can be presented as follows:

$$
\Sigma_u = \text{cov}(N \cdot \hat{\theta}) = N \cdot \text{cov}(\theta) \cdot N^T = N \cdot \Sigma_\theta \cdot N^T.
$$

According to Eqs. $(S-1.9) - (S-1.11)$, $\Sigma_u$ can be computed as follows:

$$
\Sigma_u = N \cdot (H_{SSR})^{-1} \cdot N^T = N \cdot (H_{SSR})^{-1} \cdot H_{SSR}^{-1} \cdot (H_{SSR})^{-1} \cdot N^T =
\Sigma_u = N \cdot [H_{SSR}(\theta)]^{-1} \cdot [J_i(\hat{\theta})]^T \cdot (\Sigma_\theta)^{-1} \cdot J_i(\hat{\theta}) \cdot [H_{SSR}(\theta)]^{-1} \cdot N^T.
$$

Thus, the form where the linearized approximation of the flux variances,

$$(\sigma_{lin}^u)_i = \sqrt{(\Sigma_\theta)_{ii}}, \quad i = 1, 2, ..., n,$$

is linked with the measurement variances,

$$(\sigma_{mea}^u)_j = \sqrt{(\Sigma_\theta)_{jj}}, \quad j = 1, 2, ..., w,$$

has been obtained. A set of linearized statistics parameters can be computed using this equation. In particular, the linearized approximation of the flux confidence intervals at the confidence level of $\gamma$, $CI_{\gamma}^{lin}(u_i)$, can be estimated [S-4, S-17] through the $(\sigma_{lin}^u)_i$ variances as follows:

$$
CI_{\gamma}^{lin}(u_i) = \left[\hat{u}_i - \sqrt{\gamma^2(1) \cdot (\sigma_{lin}^u)_i}, \hat{u}_i + \sqrt{\gamma^2(1) \cdot (\sigma_{lin}^u)_i}\right].
$$

These intervals must be additionally constrained according to Eq. $(S-1.2)$. 
Flux parameters that are characterized by the local linearized statistics could be used for optimizing the experimental design according to the method supposed by Möllney et al. [S-10]. In this study, a comparison of different experimental designs was provided by the $D$-criterion [S-26], which means evaluation of the squared volumes of the confidence ellipsoids with the size of the axis equaled to $CI_{\text{lin}}^\text{est}(u_i)$ for a given confidence level. According to [S-10], the $D$-factor, proportional (up to a constant that depends on confidence level $\gamma$) to the squared volume of the $p$-dimensional ellipsoid for the variances of the optimized free fluxes, is given by the following equation:

$$D(m^\text{input}, m^\text{mea}, \sigma^\text{mea}, \hat{\theta}) = \det \text{cov} (m^\text{input}, m^\text{mea}, \sigma^\text{mea}, \hat{\theta}) = \det \Sigma_{\theta} = \det [H_{\text{SSR}}(\hat{\theta})]^{-1}, \quad (S-1.5.14)$$

where the last equality, which is based on Eq. (S-1.5.10), is used for the $D$-factor computation in OpenFLUX2. Several properties of Hessian matrix (like positive definite, condition number and etc.) are additionally controlled to provide reliability of $D$-factor calculation. The $D$-factor must be minimized with respect to the space of all feasible input tracers, $m^\text{input}$, if the minimum ellipsoid volume for estimated free flux variances is the aim of the experimental design optimization as follows:

$$D_{\text{opt}}(m^\text{input}, m^\text{mea}, \sigma^\text{mea}, \hat{\theta}) = \min_{m^\text{mea-feasible}} \det [H_{\text{SSR}}(\hat{\theta})]^{-1}. \quad (S-1.5.15)$$

Simultaneously, if the minimization of the targeted flux variance equaled to $\sqrt{(\Sigma_{\theta})_{ii}}$ is desired, then it is necessary to solve the following problem:

$$\min_{m^\text{mea-feasible}} \left( \sigma^\theta_{ii} \right) = \min_{m^\text{mea-feasible}} \sqrt{\left( H_{\text{SSR}}(\hat{\theta}) \right)_{ii}^{-1}}. \quad (S-1.5.16)$$

Another field where the flux parameters evaluated by local linearized statistics play an important role in $^{13}$C-MFA is in computing and applying the so called “contribution matrix”, which, in fact, is the squared value of the
Pearson’s linear correlation coefficients between fluxes and measurements calculated based on a linear approximation of the nonlinear vector-function $f(x^{\text{input}}, \theta)$ using Taylor’s expansion. According to [S-4], Eq. (S-1.5.12) is used to construct the contribution matrix, $\text{CM}$, $\text{dim}(\text{CM}) = (n \times w)$, with the $(\text{CM})_{h,j}$ elements manifesting the relative importance of the variance of the $j$-th measurement to the approximated variance of the $i$-th flux as follows:

$$(\text{CM})_{h,j} = \frac{\left( N \cdot [H_{\text{SSR}}(\theta)]^{-1} \cdot [J_t(\theta)]^T \right)_{h,j}^2}{(\Sigma_u)_{h,j} \cdot (\Sigma_u)_{j,j}} = \frac{\left( N \cdot [H_{\text{SSR}}(\theta)]^{-1} \cdot [J_t(\theta)]^T \right)_{i,j}^2}{(\sigma^u_{\text{lin}}) \cdot (\sigma^u_{\text{mea}})_{j,j}},$$

(S-1.5.17)

where the symmetry of the Hessian matrix, i.e., $H_{\text{SSR}} = (H_{\text{SSR}})^T$ (and thus, $(H_{\text{SSR}})^{-1} = ((H_{\text{SSR}})^{-1})^T$) is used in the equation derivation process. Thus, the $(\text{CM})_{h,j}$ elements with large values indicate that the $j$-th measurement is important for evaluating the $i$-th flux. Furthermore, the sum of elements in an arbitrary row of the contribution matrix equals one (see SF-1 Appendix 3). In turn, the sum of elements of an arbitrary $j$-th column of the $\text{CM}$ matrix indicates the relative importance of the corresponding $j$-th measurement. Consequently, the large value of the sum column elements indicates the measurement that significantly contributes to the solution of the flux estimation problem. Usually [S-4], the tracer experiment is considered inadequately designed when a significant number of estimated fluxes depends on a small number of measurements.

Computing the contribution matrix according to Eq. (S-1.5.17) was incorporated into OpenFLUX2.

**SF-1.6. Nonlinear search of flux confidence intervals**

The nonlinear-based approach developed by Antoniewicz et al. [S-4] was implemented in OpenFLUX(2) to search for an estimation of the flux confidence
intervals, $CI_{\gamma}^{n-lin}(u_i)$, which is more accurate than the linear approximation, $CI_{\gamma}^{lin}(u_i)$. This approach is based on the assumption that provided measurement’s errors data for the flux model used in SLEs (or in PLEs (Exp. (S-1.4.1))) follows normal distribution and correctly weights model’s residuals. Therefore, at a convergence point, $\mathbf{u}(\hat{\theta}) = (u_1(\hat{\theta}), u_2(\hat{\theta}), ..., u_n(\hat{\theta}))^T$, the value $\xi_{\Xi} = \min_{\theta \in \mathbb{R}^p}(\Xi(\theta))$ could be treated as a random variable, which follows a $\chi^2$-distribution with $(W - p)$ degrees of freedom. For the $u_i(\theta)$ flux, let $\psi_i(v) = \min_{\theta \in \mathbb{R}^p}(\Xi(\theta; u_i(\theta) = v))$ denote the optimal value of the $\Xi(\theta)$ function when the $u_i(\theta)$ flux is fixed at the point, where $u_i(\theta) = v$. Assuming the absence of mathematical singularity for the $\Xi(\theta)$ function around a point of convergence, the value of the function $\psi_i(v)$ for fixed $v$ can be considered as a random variable that follows $\chi^2$-distribution with $(W - p - 1)$ degrees of freedom. Due to the additivity of the $\chi^2$-distribution, the non-negative difference $(\psi_i(v) - \xi_{\Xi})$ follows a $\chi^2$-distribution with one degree of freedom. According to Antoniewicz et al. [S-4], an approximate $\gamma \times 100\%$ confidence interval for $u_i$, $CI_{\gamma}^{n-lin}(u_i)$, is given by the flux values for which the following inequality is true:

$$\psi_i(v) < \xi_{\Xi} + \chi^2_{\gamma}(1). \quad (S-1.6.1)$$

In OpenFLUX(2), the lower and upper bounds, $LB_{\gamma}^{n-lin}(u_i)$ and $UB_{\gamma}^{n-lin}(u_i)$ of the corresponding confidence interval $CI_{\gamma}^{n-lin}(u_i)$ are obtained as a solution by MATLAB’s FMINCON of the following non-linear optimization problems:
\[
\begin{align*}
LB_{\gamma}^{n-lin}(u_i) &= \min_{\theta} u_i(\theta), \\
\Xi(\theta) - \xi_\Xi < \chi^2_\gamma(1), \\
\chi^2_\gamma(1) - (\Xi(\theta) - \xi_\Xi) < \beta, \\
u_i(\theta) \leq u_i(\bar{\theta}) & \quad (S-1.6.2) \\
\{ \theta \in \mathbb{R}^p \}
\end{align*}
\]

where \( \beta \) – a user-provided constraint tolerance, and the optimization procedure is started from the convergence point, \( u_i(\bar{\theta}) \), for each \( CI_{\gamma}^{n-lin}(u_i) \) bound.

Thus in OpenFLUX(2), the \( CI_{\gamma}^{n-lin}(u_i) \) intervals determined at the confidence level of \( \gamma \) are presented as follows:

\[
CI_{\gamma}^{n-lin}(u_i) = [LB_{\gamma}^{n-lin}(u_i), UB_{\gamma}^{n-lin}(u_i)].
\]  

(S-1.6.3)

It should be noted, that, despite the fact that solution to the optimization problems \( S-1.6.2 \) gives the correct minimal and maximal flux values, which satisfy inequality \( S-1.6.1 \), computational approach proposed in OpenFLUX(2) is different from the algorithm proposed by Antoniewicz et al. \( S-4 \). Nevertheless, these computational approaches are equivalent in case, when \( \psi_i(v) \) is a strictly monotone increasing function.

Since the lower and upper bounds of \( CI_{\gamma}^{n-lin} \) are obtained as a result of fitting procedure, it is important to perform detailed analysis of the optimization results. Original implementation of the non-linear search algorithm in OpenFLUX detects simple cases, when acquired value for the \( CI_{\gamma}^{n-lin} \) bounds may be unreliable, but does not provide tools for comprehensive analysis of the quality of acquired approximation of the respective \( CI_{\gamma}^{n-lin} \), and control of error propagation in optimization procedure.

It was detected, that in some cases significant bias was introduced in \( CI_{\gamma}^{n-lin} \) estimation, which made it impossible to compare confidence intervals
for different $\gamma$ (e.g., for certain models it could be observed, that estimated $CI_{0.95}^{n-lin}$ were significantly narrower than $CI_{0.68}^{n-lin}$). Several examples of these “mistaken” evaluations usually detected for the minor part of the tested fluxes, were described in the item 2.4 of Results.

One of the reasons for poor $CI_{\gamma}^{n-lin}$ estimation results was the fact that original OpenFLUX non-linear confidence intervals estimation procedure did not take into consideration the first-order optimality values calculated during FMINCON execution. These values represent fulfillment of necessary optimality conditions (Karush-Kuhn-Tucker conditions see, e.g., [S-21]) for constraint optimization problems. Thus, for certain metabolic models and function’s termination tolerances, it led to termination of the optimization procedure before it reached the lower or upper bounds of the corresponding $CI_{\gamma}^{n-lin}(u_i)$. Original OpenFLUX software could incorrectly consider that optimization had converged in such cases. Thus, it is necessary to perform additional evaluation of precision of $CI_{\gamma}^{n-lin}$ obtained by non-linear search algorithm.

Furthermore, adjustment of optimization parameters (e.g., function termination tolerance and constraint tolerance) can lead to significant improvement of first-order optimality values and consequently precision of $CI_{\gamma}^{n-lin}$ determined by non-linear confidence estimation algorithm implemented in OpenFLUX.

Despite the fact, that current version of OpenFLUX2 still utilizes original implementation of the $CI_{\gamma}^{n-lin}$ estimation algorithms, several modifications targeted to improve precision of computations, are now in the process of development.
One of the main advantages of the $CI_{\gamma}^{n-lin}$ evaluation is its computation speed that is significantly higher than for more precise, but rather slow Monte Carlo-based approach (see, the item 2.4 of Results, and SF-1.7 below). It means, that even the current, not perfect non-linear search algorithm implemented in OpenFLUX(2) could be efficiently used for the fast preliminary evaluation of the flux confidence intervals according to Eq. (S-1.6.3). For the most part of fluxes, these evaluations are close to the precise parameters that could be determined later by Monte Carlo-based approach.

Using the computed $CI_{\gamma}^{n-lin}(u_i)$ parameters, it is possible to assign the value of the $i$-th “flux standard deviation” or “flux precision” [S-18], $\sigma_{\gamma}^{n-lin}(\gamma)$, instead the linear approximations $\sigma_{\gamma}^{lin}$ (see, SF-1.5), by the following way:

$$\sigma_{\gamma}^{n-lin}(u_i) = \frac{UB_{\gamma}^{n-lin}(u_i) - LB_{\gamma}^{n-lin}(u_i)}{2 \cdot \sqrt{\chi^2_{\gamma}(1)}}$$  \hspace{1cm} (S-1.6.4)

formally comparing Eqs. (S-1.5.13) and (S-1.6.3). Eq. (S-1.6.4) can be simplified [S-18] for the $\gamma = 0.95$ (95% confidence interval) to the following equation:

$$\sigma_{0.95}^{n-lin}(u_i) = \frac{1}{2 \cdot \sqrt{3.841}} \cdot (UB_{0.95}^{n-lin}(u_i) - LB_{0.95}^{n-lin}(u_i)) \approx \frac{1}{4} \cdot (UB_{0.95}^{n-lin}(u_i) - LB_{0.95}^{n-lin}(u_i)).$$ \hspace{1cm} (S-1.6.5)

Namely the value of $\sigma_{0.95}^{n-lin}(u_i)$ have been earlier used for evaluation the achieved flux resolution [S-18, 27]. Our proposal is to introduce the function $\eta_{\gamma}^{n-lin}(u_i(\hat{\theta}), \beta)$, which is assigned as “the normalized flux precision at a determined confidence level of $\gamma$” or “normalized flux precision”, which must be computed in the following manner:
\[ \eta^\text{n-lin}_y(u_i(\hat{\theta}), \beta) = \begin{cases} 1 - \frac{CI^\text{n-lin}_y(u_i)}{u_i(\hat{\theta}) + \beta \cdot \text{max } \text{V}_{\text{mea}}}, & \text{when } CI^\text{n-lin}_y(u_i) \leq u_i(\hat{\theta}) + \beta \cdot \text{max } \text{V}_{\text{mea}} \\ 0, & \text{when } CI^\text{n-lin}_y(u_i) > u_i(\hat{\theta}) + \beta \cdot \text{max } \text{V}_{\text{mea}} \end{cases} \] (S-1.6.6)

where \( \beta \) is a scaling parameter and \( \text{max } \text{V}_{\text{mea}} \) is the maximal measured efflux, which, usually, corresponds to input carbon flux. It should be noted that the \( u_i(\hat{\theta}) \) flux value is the best available estimation computed from the SLE- or PLE-based \(^{13}\text{C}-\text{MFA} \), which approximates the unknown true flux value. In the present study, the known values of the true fluxes were used for estimation of flux precision efficiencies achieved due to analysis of different LEs performed \textit{in silico}. According to Eq. (S-1.6.6), in the case of rather small or zero values of flux parameters their estimation of flux precision is evaluated relative to the \( \text{max } \text{V}_{\text{mea}} \), as some characteristic value. A variable scaling parameter \( \beta \) provides flexibility in determination of what is the “slow” fluxes, depending on the case context. The scaling parameter \( \beta \) had been set to 0.1 at the present study.

According to its definition, the \( \eta^\text{n-lin}_y(u_i(\hat{\theta}), \beta) \) function at each fixed \( \beta \) parameter is close to “1” for precisely estimated flux \( u_i(\hat{\theta}) \) (with narrow confidence interval) and is close to “0” for poorly determined flux \( u_i(\hat{\theta}) \).

Analogously, the normalized flux precision functions can be defined using the flux confidence intervals determined using the linearized approximation of the flux variances (see Eq. (S-1.5.13)) or the Monte Carlo-based approach (see \textit{SF-1.7}). The corresponding superscript, \( \text{lin}, \text{n-lin}, \text{MC-1} \) or \( \text{MC-2} \) in the definition of the flux precision function, e.g., \( \eta^\text{MC-1}_y(u_i(\hat{\theta}), \beta) \) or \( \eta^\text{lin}_y(u_i(\hat{\theta}), \beta) \), directly indicates the method used for computing the flux confidence intervals.
**SF-1.7. Determination of flux confidence intervals using the Monte Carlo-based approach**

Usually in the present study, the flux confidence intervals were determined using the Monte Carlo-based approach included previously in OpenFLUX [S-6], but substantively modified in the final design of OpenFLUX2 to allow fine-tuning of procedures parameters and add convergence control.

According to [S-16], the Monte Carlo approach of flux confidence interval search was based on discrete approximation of optimized flux parameter distributions obtained in $L$ multi-trials when the measurement’s data for each trial were artificially generated by applying normally distributed random errors to initially obtained experimental measurements data. The number of trials, $L$, is one of the important parameters of Monte Carlo-based procedure, and it has to be optimally chosen ($L = L_{\text{MAX}}$) for flux confidence intervals, $CI_r^{\text{MC}}$, precise determination in a reasonable computation time (see, below). The data variations were performed by corrupting the MIDs and the effluxes measured with the corresponding deviations, $\sigma_{\text{MID}}^{\text{mea}}$ and $\sigma_{\text{eff}}^{\text{mea}}$. Random errors, which followed $N(0, \Sigma_{\epsilon})$ distribution (with the relation in accordance to Eq. (S–1.3.3) between $\Sigma_{\epsilon}$, $\sigma_{\text{MID}}^{\text{mea}}$, and $\sigma_{\text{eff}}^{\text{mea}}$) were generated using the statistics Toolbox of MATLAB. Then, the generated $L$ data sets, $m_j^{\text{mea}}$, $j = 1, 2, ..., L$, were used as new “experimental” data in corresponding $L$ constrained NLLSPs (as in case of Eq. (S–1.3.8)) to generate the following set of optimized flux estimations using the fitting-based quasi-Newton SQP method (see **SF-1.3.**):

$$\textbf{U}(\hat{\Theta}_L) = (\textbf{u}(\hat{\Theta}_1), \textbf{u}(\hat{\Theta}_2), ..., \textbf{u}(\hat{\Theta}_L), ..., \textbf{u}(\hat{\Theta}_L)),$$

(S–1.7.1)
where \( \hat{\theta}_j = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_j) \) and \( u(\hat{\theta}_j) = (u_1(\hat{\theta}_j), u_2(\hat{\theta}_j), \ldots, u_{\ell}(\hat{\theta}_j)) \), \( j = 1, 2, \ldots, L \). Two different methods were implemented in OpenFLUX(2) to decrease computational time of the optimization process of each trial in Monte Carlo-based search of \( C_{\gamma}^{MC} \) in comparison with the time needed to acquire optimized flux parameters as a result of the search for the solution of the initial constrained NLLSP.

The first method consists of a single run of the optimization procedure for each trial, that uses optimized free fluxes \( \hat{\theta} \), which were obtained for the uncorrupted measurement data, as a starting point. This method, “single run per trial”, results in fast performance but tends to determine the local minimum instead of the global minimum.

The second method incorporates a simple technique, which attempts to determine the global minimum of the \( \Xi(\theta) \) function. Instead of a single run in each trial, this technique consists of sequential multiple optimization runs from random starting points \( \theta \in \mathbb{R}^\ell \), that additionally satisfy the inequality (S–1.2.1). As described previously (see SF-1.3.), each run proceeds through step-wise iterative changes of free flux values until reaching the minimum of the target objective function with the tunable termination criterion, TT – termination tolerance, which is by default equal to \( 1 \times 10^{-4} \) in OpenFLUX2 (it is a tunable parameter in Monte Carlo search of flux confidence intervals). In each trial, the set of the best found local minimum is stored during optimization. If the new run results in a better optimized value \( \xi' = \Xi(\theta') \) of \( \Xi(\theta) \), then the set is emptied, and the new value \( \xi' \) and the corresponding set of flux values \( u(\theta') \) is stored. If the value \( \xi' \) of the current run is close to the currently stored best
value \( \hat{\xi}_e = \Xi(\hat{\Theta}) \), then this run solution is added to the set. Formally, the distance between the corresponding values (proximity) is defined by the following expression:

\[
\rho\left(\hat{\xi}_e', \hat{\xi}_e\right) = \left| \frac{\hat{\xi}_e - \hat{\xi}_e'}{\hat{\xi}_e} \right|, \tag{S-1.7.2}
\]

Thus, the current optimized value \( \xi'_e \) is added to the set when \( \rho\left(\xi'_e, \xi_e\right) \leq \varepsilon \), where \( \varepsilon \) is a predefined proximity threshold (it is a tunable parameter in OpenFLUX2). The value \( \varepsilon = 0.05 \) has been initially used as a default in OpenFLUX(2), but it is highly recommended to decrease the default value up to \( \varepsilon = 1 \times 10^{-4} \) or below (see, Figure SF-2.4), to improve precision of \( C_{\gamma}^{\text{MC}} \) determination with the help of OpenFLUX2. Notably, runs that encounter any problems during optimization step (e.g., the optimization procedure stops before reaching a prescribed convergence threshold) are completely ignored. Accordingly, the reiterative run solutions are accumulated. The procedure terminates when the size of the accumulated set exceeds a certain threshold value \( N_{\text{AS}} \) (it is a tunable parameter, and by default \( N_{\text{AS}} = 3 \) in OpenFLUX2). Additionally, it is possible to constrain the number of runs in one trial. In this case, optimization procedure in each trial will terminate after performing the prescribed number of runs \( K_{\text{NR}} \) (it is a tunable parameter, and by default \( K_{\text{NR}} = 50 \) in OpenFLUX2), when the set of accumulated run solutions is not filled completely, and the parameters corresponding to the best stored run would be considered as optimized for this trial. This mandatory termination leads to faster performance but can potentially demonstrate poorer results in comparison with the case when the number of run per trial is not restricted.
Certainly, optimization procedure that was performed according to the second strategy, “multiple runs per trial”, results in slower performance, but increases the chances of finding the global minimum for each out of \( L \) trials and, therefore, better estimates the flux confidence intervals. Namely, this strategy is used as a default approach for Monte Carlo procedure in OpenFLUX2, but the alternative, “one run per trial”, method could be used according to the user’s choice. Furthermore, in OpenFLUX2, it is possible to fine-tune the TT, \( \varepsilon \), \( N_{AS} \) and \( K_{NR} \) parameters of the above-described procedure to govern solution accuracy and computation time of each trial. For example, in case of unrestricted number of runs per trial, values for the proximity threshold \( \varepsilon \) and the size of the accumulated run solutions set \( N_{AS} \) can be decreased or increased in comparison with the default values, correspondingly, to improve precision of the \( CI_{MC}^{\gamma} \) determination. Usually, these improvements resulted in narrowing of the \( CI_{MC}^{\gamma} \) width, that, in turn, resulted in decrease of the upper bound, and in increase of the lower bound of the target flux confidence interval. Computed \( CI_{MC}^{\gamma} \)'s estimations for one of the fluxes as a function of different values of tunable parameters, which determine the accuracy of the constrained NLLSP solution in each trial of Monte Carlo-based approach, are presented in the Figure SF-2.4.

Therefore, one of the above-described strategies is used iteratively to generate a set of independently obtained flux distributions of size \( L \) (Eq. (S-1.7.1)). To determine the flux confidence intervals on the basis of these distributions, two different approaches are implemented in OpenFLUX2. According to the first approach, “discarding” strategy, for each \( i, i = 1, ..., n \), the set \( U_i(\hat{\theta}_L) = \{u_i(\hat{\theta}_1), u_i(\hat{\theta}_2), ..., u_i(\hat{\theta}_L)\} \) of flux estimated values is first sorted
in ascending order. Then, the $\gamma \times 100\%$ (e.g., 95%, 80 or 68%) confidence interval on the basis of $L$ trials, $CI_{\gamma}^{MC-1}(U_i(\hat{\Theta}_L))$, is obtained by discarding the top $\frac{1-\gamma}{2} \times 100\%$ (e.g., 2.5%, 10 or 16%) and the bottom $\frac{1-\gamma}{2} \times 100\%$ of the flux estimates $u_i(\hat{\Theta}_L), j = 1, \ldots, L$, which results (see, for example, [S-4]) in determination of upper and lower bounds $UB_{\gamma}^{MC-1}(u_i; L)$ and $LB_{\gamma}^{MC-1}(u_i; L)$ respectively, for each $u_i(\hat{\Theta}_L), i = 1, 2, \ldots, n$ at a confidence level of $\gamma$ (see, Figure 4 in the main text for clarity).

Alternatively [S-6, S-16, S-17], the “mean-varianced” strategy could be used, where $i$-th flux mean, $\mu_i^{MC-2}(\hat{\Theta}_L)=\frac{1}{L} \sum_{j=1}^{L} u_i(\hat{\Theta}_j)$, and the unbiased estimator of the variance, $\sigma_i^{MC-2}(\hat{\Theta}_L)=\frac{1}{(L-1)} \left| \sum_{j=1}^{L} (u_i(\hat{\Theta}_j) - \mu_i^{MC-2}(\hat{\Theta}_L))^2 \right|$, are calculated to generate the confidence intervals at the confidence level of $\gamma$ on the basis of $L$ trials, $CI_{\gamma}^{MC-2}(U_i(\hat{\Theta}_L))$, which, finally, must be additionally constrained by the range $[lb_i, ub_i]$ of the corresponding feasible values for the $i$-th flux as follows:

$$CI_{\gamma}^{MC-2}(U_i(\hat{\Theta}_L))=\left[ \max \{lb_j, \mu_i^{MC-2}(\hat{\Theta}_L) - \Delta_i(\hat{\Theta}_L)\}, \min \{\mu_i^{MC-2}(\hat{\Theta}_L) + \Delta_i(\hat{\Theta}_L), ub_j\} \right], \quad (S-1.7.3)$$

where $\Delta_i(\hat{\Theta}_L)=\sigma_i^{MC-2}(\hat{\Theta}_L) \sqrt{\chi^2_{\gamma}(1)}$.

Finally, the Monte Carlo-based confidence intervals computed by “discarding”, $CI_{\gamma}^{MC-1}(u_i)$, or “mean-varianced”, $CI_{\gamma}^{MC-2}(u_i)$, strategies have to be determined as:

$$CI_{\gamma}^{MC-1}(u_i) = \lim_{L \to \infty} CI_{\gamma}^{MC-1}(U_i(\hat{\Theta}_L)), \quad CI_{\gamma}^{MC-2}(u_i) = \lim_{L \to \infty} CI_{\gamma}^{MC-2}(U_i(\hat{\Theta}_L)) \quad \text{(S-1.7.4)}$$

i.e., the estimated values of upper and lower bounds of $CI_{\gamma}^{MC}(u_i)$ have to converge in case of significant increase of the total number of trials. The special
ONLINE procedure has been implemented in OpenFLUX2 to control the number of optimization trials that is performed during estimation of target flux $\hat{C}_{\gamma}^{MC}(u_i)$ bound.

The designed control procedure is based on the assumption that increase in the number of trials results in improvement of approximation precision. That, in turn, has to result in convergence of $\hat{C}_{\gamma}^{MC}(u_i)$ bounds. In OpenFLUX2, the method of “sliding control” has been applied for evaluation of this convergence. Let, $M$ be a preset number that characterized a sliding control “window” size. First, for each set $U_i(\tilde{\Theta}_{L-k})$ (see, Eq. $(S - 1.7.1)$), $k = 0, 1, \ldots, (M - 1)$, and $i$-th flux, $i = 1, 2, \ldots, n$, at confidence level of $\gamma$ the bounds of $\hat{C}_{\gamma}^{MC}(u_i)$ are estimated, using, for example, the “discarding” strategy:

$$\hat{C}_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k})) = [LB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k})), UB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))].$$

Then an absolute spread in the estimated lower (and upper) bounds of flux confidence intervals can be finally calculated as:

$$\Delta_{LB}^{MC-1}(\gamma, i, \tilde{\Theta}_{L}, M) = \max_{0 \leq k \leq M-1} LB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))-\min_{0 \leq k \leq M-1} LB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k})), \quad (S - 1.7.6)$$

$$\Delta_{UB}^{MC-1}(\gamma, i, \tilde{\Theta}_{L}, M) = \max_{0 \leq k \leq M-1} UB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))-\min_{0 \leq k \leq M-1} UB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k})).$$

Consequently, values of a relative spread for all estimated bounds can be determined as:

$$\delta_{LB}^{MC-1}(\gamma, i, \tilde{\Theta}_{L}, M) = \frac{2 \times \Delta_{LB}^{MC-1}(\gamma, i, \tilde{\Theta}_{L}, M)}{\max_{0 \leq k \leq M-1} LB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))+\min_{0 \leq k \leq M-1} LB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))}, \quad (S - 1.7.7)$$

$$\delta_{UB}^{MC-1}(\gamma, i, \tilde{\Theta}_{L}, M) = \frac{2 \times \Delta_{UB}^{MC-1}(\gamma, i, \tilde{\Theta}_{L}, M)}{\max_{0 \leq k \leq M-1} UB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))+\min_{0 \leq k \leq M-1} UB_{\gamma}^{MC-1}(U_i(\tilde{\Theta}_{L-k}))}. \quad (S - 1.7.7)$$
If the obtained relative spread does not exceed the predetermined threshold for the targeted flux, it could be considered that the corresponding bound of $CI_{\gamma}^{MC}(u_i)$ converged (see, Figure 5 for clarity).

OpenFLUX2 allows fine tuning of the Monte Carlo-based procedure in such a fashion, that the convergence for all $CI_{\gamma}^{MC}(u_i)$ bounds results in termination of calculations, even if the predetermined maximal number of trials, $L_{MAX}$, has not been achieved yet. This option can help to optimize the number of performed trials, and, in several cases, significantly decrease computation time for Monte Carlo-based approach.

OpenFLUX2 allows to fine-tune sliding control “window” to allow the user to balance between computation time and precision of $CI_{\gamma}^{MC}(u_i)$ determination. Indeed, a user can tune different parameters of “window” (e.g., size, deepness, frequency), modify the value of predetermined threshold for $CI_{\gamma}^{MC}$ bounds. To save computation time, the user can preset the frequency of a sliding control tests (e.g., once per 40 performed trials). Moreover, OpenFLUX2 allows to analyze using the described sliding control “window” strategy estimations of $CI_{\gamma}^{MC}$, which were earlier obtained by Monte Carlo-based approach. The auxiliary visualization tools were implemented in OpenFLUX2 that allow plotting of optimized flux distributions, as well as displaying convergence plots of $CI_{\gamma}^{MC}$ bounds as a function of the number of performed trials (see, Figure 5, Figure SF-2.4, as examples).

That is to say, an automatic convergence control is implemented at the stage of $CI_{\gamma}^{MC}$ bound computation in OpenFLUX2: the several tunable parameters are introduced for evaluation of achieved convergence of the values
for the bounds as a function of increased $L$. Moreover, if several $CI_{\gamma}^{MC}(u_i)$ bounds for certain fluxes has not reached convergence for the chosen set of control parameters, then the calculation is stopped at the $L=L_{\text{max}}$, and convergence plots of the corresponding $CI_{\gamma}^{MC}$ bounds as well as corresponding diagrams of optimized fluxes can be visualized on the screen. In this case, user has to make the final decision concerning $CI_{\gamma}^{MC}$ parameters for these questionable fluxes.

The mentioned automatic convergence control can be applied for several targeted fluxes (e.g., for all free fluxes, or for all fluxes) or for an arbitrary function of free fluxes (net fluxes, flux ratios, etc.). Certainly, the number of trials necessary to achieve convergence for flux confidence intervals (and proportionally the total computation time) can be different depending on the chosen target. If the convergence for all free fluxes $CI_{\gamma}^{MC}$ bounds has been already achieved, then evaluation of $CI_{\gamma}^{MC}$ for residual dependent fluxes, or for any other functions of free fluxes can be finally continued with the help of MATLAB’s FMINCON function. But according to our experience (frankly speaking, not very significant), it is possible to achieve rather satisfactory results concerning evaluation of $CI_{\gamma}^{MC}$ for the most of questionable fluxes after visualization of the plots when no more than 1,000 – 2,000 trials of Monte Carlo-based approach were performed, without additional exploiting of FMINCON.

It could be added, that when the $U_i(\hat{\Theta}_L), i = 1, 2, ..., n$, estimates acquired by the described Monte Carlo methods are non-symmetrically distributed, the computed $CI_{\gamma}^{MC-1}(U_i(\hat{\Theta}_L))$ values are close to the $CI_{\gamma}^{\text{lin}}(u_i)$ obtained using the
nonlinear-based approach (1.6) [S-4]. Thus, it seems preferable to use the estimations of $\mathcal{C}_i^{MC^{-1}}(u_i)$ obtained by the “discarding” strategy in the Monte Carlo-based approach; however, the evaluations provided according to Eq. (S–1.7.3) are also retained in OpenFLUX2.
**SF-1 Appendix 1. The special form of the null-space matrix in OpenFLUX(2)**

Suppose that the metabolic system consisted of \( t \) bi-directional (reversible) reactions and \( s = n - 2t \) unidirectional (irreversible) reactions. It is possible to reorder the variables in the initial flux vector \( \mathbf{v} \) for generating the following vector:

\[
\mathbf{v'} = \begin{pmatrix} \mathbf{v^\rightarrow} \\
\mathbf{v^{irr}} \\
\mathbf{v^\leftarrow} \end{pmatrix},
\]

(A1-1)

where \( \mathbf{v^\rightarrow} \) and \( \mathbf{v^\leftarrow} \) are \( t \)-dimensional column vectors consisting of all fluxes corresponded to forward and reverse bi-directional reactions, and \( \mathbf{v^{irr}} \) is the \( s \) -dimensional column vector corresponding to all irreversible reactions.

The reordering of variables in the flux vector automatically leads to the corresponding reordering of the columns in the \( \mathbf{S} \) matrix. Considering Eq. (S-1.1.1), the \( \mathbf{S} \) corresponding to the vector \( \mathbf{v'} \) could be presented in the following form:

\[
\mathbf{S} = \begin{pmatrix} \mathbf{S^\rightarrow} & \mathbf{S^{irr}} & -\mathbf{S^\leftarrow} \end{pmatrix},
\]

(A1-2)

where \( \mathbf{S^\rightarrow} \) and \( \mathbf{S^{irr}} \) are the matrixes consisting of the columns of \( \mathbf{S} \) that correspond to \( \mathbf{v^\rightarrow} \) and \( \mathbf{v^{irr}} \), respectively. Because \( \text{rank}(\mathbf{S}) = \text{rank}(\mathbf{\tilde{S}}) = r \), Eq. (S-1.1.1) could be converted to the equivalent equation:

\[
\mathbf{\tilde{S}} \cdot \mathbf{v'} = 0.
\]

(A1-3)

Let \( \mathbf{v^{\rightarrow,net}} \) be a new net forward flux vector that is determined as follows:

\[
\mathbf{v^{\rightarrow,net}} = \mathbf{v^\rightarrow} - \mathbf{v^\leftarrow},
\]

(A1-4)

where, in contrast to other flux vectors that always have their entries \( \geq 0 \), \( \mathbf{v^{net}} \) from \( \mathbf{v^{\rightarrow,net}} = (v_1^{net}, v_2^{net}, ..., v_t^{net})^T \) could be positive or negative depending on the
scalar values of corresponding $v_i^+$ forward and $v_i^-$ reverse fluxes of $i$-th, $1 \leq i \leq t$ bi-directional reaction. Then, let a new flux vector, $v''$, be introduced according to the following equation:

$$v'' = \begin{pmatrix} v_i^{+,net} \\ v_{irr} \\ v_i^e \end{pmatrix}. \quad (A1-5)$$

The following relation can be detected between $v'$ and $v''$:

$$v' = \begin{pmatrix} v_i^+ \\ v_{irr} \\ v_i^- \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_i^{+,net} \\ v_{irr} \\ v_i^e \end{pmatrix} = A \cdot v'', \quad (A1-6)$$

where matrix $A$ is assigned as follows:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (A1-7)$$

Let $\tilde{S}'$ and $v''$ be $\tilde{S}' = \begin{pmatrix} S^{rev} \\ S^{irr} \end{pmatrix}$ and $v'' = \begin{pmatrix} v_i^{+,net} \\ v_{irr} \end{pmatrix}$, respectively. Then, in accordance with $(A1-2)$ and $(A1-6)$, the following equalities must be accomplished:

$$\tilde{S} \cdot v' = \tilde{S} \cdot (A \cdot v'') = (\tilde{S} \cdot A) \cdot v'' = \begin{pmatrix} S^{rev} \\ S^{irr} \\ -S^{rev} \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot v'' =$$

$$= \begin{pmatrix} S^{rev} \\ S^{irr} \\ 0 \end{pmatrix} \begin{pmatrix} v_i^{+,net} \\ v_{irr} \\ v_i^e \end{pmatrix} = \tilde{S}' \cdot v''. \quad (A1-8)$$

As a result, the mass balance equation $(A1-3)$ is converted to the following equivalent equation:

$$\tilde{S}' \cdot v'' = 0, \quad (A1-9)$$
in relation to the new flux vector, $v''$. Because the $\tilde{S}'$ matrix has been obtained from the $\tilde{S}$ matrix due to the elimination of the linearly dependent columns, the ranks of the mentioned matrixes are equal.

Notably, if some linearly dependent columns are present in the first $t$ columns of the $\tilde{S}'$ matrix corresponding to $v^{-x,net}$, then the assumed metabolic model includes the cycle consisting of only bi-directional reactions. Certainly, this cycle could be presented as two cycles consisting of only direct and of only reverse fluxes of these bi-directional reactions because the parameters of the fluxes for the cycle consisting of the reverse fluxes potentially have no natural upper limit and do not depend on any parameters of other fluxes. In this case, at least one element from $v^{-x,net}$ must be chosen as one of the free fluxes, leading to the assignment of direct and simultaneous reverse fluxes of one bi-directional reaction as free fluxes for this metabolic model. For future calculations, let us consider that there are no such cycles in the assumed model, and therefore, the first $t$ columns of $\tilde{S}'$ corresponding to $v^{-x,net}$ are linearly independent.

A new matrix, $\tilde{S}''$, could be obtained due to reordering the linearly dependent columns in the right part of the modified matrix and to the Gaussian elimination-mediated conversion of $\tilde{S}'$ to the following form:

$$\tilde{S}'' = \begin{pmatrix} I & 0 & M' \\ 0 & I & M'' \\ 0 & 0 & 0 \end{pmatrix},$$

(A1-10)

where $M'$ and $M''$ are the matrixes with dimensions $t \times (n-r-t)$ and $(r-t) \times (n-r-t)$, respectively, which have been obtained at the stage of conversion of $\tilde{S}'$ to $\tilde{S}''$. In this equation, $v''$ is converted to $\tilde{v}''$ in the following form:
\[ \mathbf{\hat{v}}^m = \begin{pmatrix} v^{\to, \text{net}} \\ v^{\text{irr}} \\ \theta^{\text{dep}} \end{pmatrix}. \]  \hfill (A1-11)

Let \( M_{\text{net}}^{\text{rev}} = -M' \) and \( M_{\text{net}}^{\text{irr}} = -M^* \), then from Eq. (A1-10), the new form of the mass balance equation could be expressed as follows:

\[ \mathbf{\tilde{S}}^*: \mathbf{\hat{v}}^m = \begin{pmatrix} I & 0 & -M_{\text{net}}^{\text{rev}} \\ 0 & I & -M_{\text{net}}^{\text{irr}} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v^{\to, \text{net}} \\ v^{\text{irr}} \\ \theta^{\text{dep}} \end{pmatrix} = 0, \]  \hfill (A1-12)

and the following equality would be true:

\[ \begin{pmatrix} v^{\to, \text{net}} \\ v^{\text{irr}} \\ \theta^{\text{dep}} \end{pmatrix} = \begin{pmatrix} M_{\text{net}}^{\text{rev}} \\ M_{\text{net}}^{\text{irr}} \\ I \end{pmatrix} \cdot \theta^{\text{irr}}. \]  \hfill (A1-13)

The \( N_{\text{net}} \) will be assigned as follows:

\[ N_{\text{net}} = \begin{pmatrix} M_{\text{net}}^{\text{rev}} \\ M_{\text{net}}^{\text{irr}} \\ I \end{pmatrix}. \]  \hfill (A1-14)

Then, the general solution of the system (A1.12) could be expressed as follows:

\[ \mathbf{\hat{v}}^m = N_{\text{net}} \cdot \theta^{\text{irr}}. \]  \hfill (A1-15)

Let \( \theta = \begin{pmatrix} \theta^{\text{irr}} \\ \mathbf{v}^{\text{dep}} \end{pmatrix} \). Then, according to Eqs. (A1-5)–(A1-7), (A1-11), and (A1-14)–(A1-15), the vector could be returned back to the initial column flux vector expressed in new variables as follows:
\[
\mathbf{u} = \begin{pmatrix} \mathbf{v}^{\rightarrow} \\ \mathbf{v}_{\text{dep}}^{\text{irr}} \\ \mathbf{0}^{\text{irr}} \\ \mathbf{v}^{\leftarrow} \end{pmatrix} = \mathbf{A} \cdot \mathbf{v}^{\rightarrow} = \mathbf{A} \cdot \begin{pmatrix} \mathbf{v}_{\text{dep}}^{\text{irr}} \\ \mathbf{0}^{\text{irr}} \\ \mathbf{v}^{\leftarrow} \end{pmatrix} = \mathbf{A} \cdot \begin{pmatrix} \mathbf{v}_{\text{net}}^{\text{irr}} \\ \mathbf{0} \\ \mathbf{I} \\ \mathbf{v}^{\leftarrow} \end{pmatrix} = \mathbf{A} \cdot \begin{pmatrix} \mathbf{N}_{\text{net}} \\ \mathbf{0} \\ \mathbf{I} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{0}^{\text{irr}} \\ \mathbf{v}^{\leftarrow} \end{pmatrix} = \mathbf{N} \cdot \mathbf{v}^{\leftarrow}.
\]

The bracketed parts of Eq. (\textit{A}1–16) correspond to Eq. (5–1.1.8) of the main text of \textit{Supplementary file 1}; therefore, the following relations could be expressed:

\[
\mathbf{v}_{\text{dep}}^{\rightarrow} = \mathbf{M}_{\text{net}}^{\text{rev}} \cdot \mathbf{0}^{\text{irr}} + \mathbf{v}^{\leftarrow}; \quad \mathbf{v}_{\text{dep}}^{\rightarrow} = \mathbf{M}_{\text{net}}^{\text{irr}} \cdot \mathbf{0}^{\text{irr}}.
\]
**SF-1 Appendix 2. Newton’s method of the constrained NLLSP solution**

In the case of NLLSP for $SSR_{t}^{SLE}$, the estimation of the optimal $\hat{\theta} \in \mathbb{R}^p$ follows an iterative solution scheme. Starting with random values of free fluxes from the constrained range, the model vector-function $f$ is evaluated at each $g$-th iteration, $\theta_g \in \mathbb{R}^p$, to obtain the values of the objective scalar-function as follows:

$$
(SSR_{t}^{SLE})_g = SSR_{t}^{SLE}(\theta_g) = \frac{1}{2} \cdot |(f(\theta_g) - m_{\text{meas}})^T \cdot (\Sigma_\epsilon(\sigma_{\text{meas}})^{-1} \cdot (f(\theta_g) - m_{\text{meas}})|, 
$$

(A2 - 1)

and to find the optimal vector of free flux changes, $\Delta \theta_g$, which leads to $SSR_{t}^{SLE}$ minimization, i.e., $\Delta (SSR_{t}^{SLE})_g = (SSR_{t}^{SLE})_{g+1} - (SSR_{t}^{SLE})_g \leq 0$. At each consequent $g$-th iteration, the $SSR_{t}^{SLE}$ function can be formally expressed around $\theta_g$ using Taylor’s expansion:

$$
SSR_{t}^{SLE}(\theta) = SSR_{t}^{SLE}(\theta_g) + \sum_{j=1}^p \frac{\partial SSR_{t}^{SLE}(\theta_g)}{\partial \theta_j} \cdot (\theta_j - (\theta_j)_g) + 
$$

$$
+ \frac{1}{2} \cdot \sum_{j=1}^p \sum_{i=1}^p \frac{\partial^2 SSR_{t}^{SLE}(\theta_g)}{\partial \theta_j \partial \theta_i} \cdot (\theta_j - (\theta_j)_g) \cdot (\theta_i - (\theta_i)_g) + O((\theta - \theta_g)^3) = 
$$

$$
= SSR_{t}^{SLE}(\theta_g) + [J_{SSR}(\theta_g)]^T \cdot (\theta - \theta_g) + \frac{1}{2} \cdot (\theta - \theta_g)^T \cdot H_{SSR}(\theta_g) \cdot (\theta - \theta_g) + O((\theta - \theta_g)^3), 
$$

(A2 - 2)

where $J_{SSR}(\theta_g)$, with $\text{dim}(J_{SSR}) = (p \times 1)$, and $H_{SSR}(\theta_g)$, with $\text{dim}(H_{SSR}) = (p \times p)$, are the Jacobian and the Hessian matrices of the $SSR_{t}^{SLE}$ scalar-function evaluated at $\theta_g \in \mathbb{R}^p$. Thus, neglecting third and higher order terms of $\Delta \theta = (\theta - \theta_g)$, $\Delta (SSR_{t}^{SLE})_g$ can be approximated by the following expression:

$$
\Delta (SSR_{t}^{SLE})_g \approx [J_{SSR}(\theta_g)]^T \cdot \Delta \theta + \frac{1}{2} \cdot \Delta \theta^T \cdot H_{SSR}(\theta_g) \cdot \Delta \theta. 
$$

(A2 - 3)

Consequently, the minimum of the $\Delta (SSR_{t}^{SLE})_g$ function is achieved when
The above-described optimization procedure continues until a predefined convergence criterion is met. Finally, the optimized values of all fluxes, \( \hat{u} = \mathbf{N} \cdot \hat{\theta} \in \mathbb{R}^n \), are calculated.

Each \( g \)-th iteration of the proposed optimization procedure requires the calculation of \( \mathbf{J}_{SSR}(\theta_g) \) and \( \mathbf{H}_{SSR}(\theta_g) \) matrixes in Eq. (A2-5). The \( \mathbf{J}_{SSR}(\theta_g) \) matrix can be analytically calculated in the following way:

\[
\mathbf{J}_{SSR}(\theta_g) = \frac{\partial \text{SSR}^\text{SLE}}{\partial \theta}(\theta_g) = \mathbf{J}_f(\theta_g) \cdot (\mathbf{E}_g) \cdot (\mathbf{f}(\theta_g) - \mathbf{m}^\text{mea}),
\]

where \( \mathbf{J}_f(\theta_g) = \frac{\partial \mathbf{f}}{\partial \theta}(\theta_g) \) is the Jacobian matrix, \( \dim(\mathbf{J}_f) = (p \times p) \), of the \( \mathbf{f} \) vector-function with respect to \( \theta \) evaluated at \( \theta_g \in \mathbb{R}^p \). Notably, this matrix is often called the matrix of the simulated measurement sensitivities or the matrix of sensitivities [S-3, S-4, S-17]. In contrast, the Hessian matrix, \( \mathbf{H}_{SSR}(\theta_g) \), cannot
be calculated directly. The following approach is usually applied for the $H_{ssr}(\theta_g)$ approximated calculation [S-4, S-8]. First, the $f$ vector-function around $\theta_g$ is formally expressed using Taylor’s expansion as follows:

$$f(\theta) = f(\theta_g) + [J_I(\theta_g)]^T \cdot (\theta - \theta_g) + O((\theta - \theta_g)^2). \quad (A2-8)$$

Using the value of $f(\theta)$ from Eq. (A2–8), neglecting the second and higher orders of $(\theta - \theta_g)$, and applying Eq. (A2–7), the following expression could be obtained for $\Delta(SSR_{SLE})_g$:

$$\Delta(SSR_{SLE})_g \approx [J_{ssr}(\theta_g)]^T \cdot \Delta \theta + \frac{1}{2} \cdot (\Delta \theta)^T \cdot [J_I(\theta_g)]^T \cdot (\Sigma_\epsilon)^{-1} \cdot J_I(\theta_g) \cdot \Delta \theta. \quad (A2-9)$$

From Eqs. (A2–3) and (A2–9) and due to the uniqueness of the form of Taylor’s series, the Hessian matrix, $H_{ssr}(\theta_g)$, can be approximated as follows:

$$H_{ssr}(\theta_g) \approx [J_I(\theta_g)]^T \cdot (\Sigma_\epsilon)^{-1} \cdot J_I(\theta_g). \quad (A2-10)$$

Equations analogous to Eq. (A2–7) and (A2–10) can be obtained for computing the $J_{ssr}(\theta_g)$ and $H_{ssr}(\theta_g)$ in Eq. (A2–6) from $f_I(\theta_g)$, $J_I(\theta_g)$, $(\Sigma_\epsilon)$, and from $(m_{\text{mea}})$, for all $i=1,2,\ldots,k$ labeling experiments from the PLE.
**SF-1 Appendix 3. Summing elements in an arbitrary row of the contribution matrix**

Let the \( C \) matrix be assigned as follows:

\[
C = N \cdot [H_{SSR}(\theta)]^{-1} \cdot [J_t(\theta)]^T,
\]

then, Eqs. \((S-1.5.12)\) and \((S-1.5.17)\) can be rewritten in the following shortened forms, respectively:

\[
\Sigma_u = C \cdot (\Sigma_x)^{-1} \cdot C^T; \quad (A3-2)
\]

\[
\{CM\}_{i,j} = \frac{C^2_{i,j}}{\{\Sigma_u\}_{i,i} \cdot \{\Sigma_x\}_{j,j}}. \quad (A3-3)
\]

Because the \( \Sigma_x \) matrix is considered diagonal (Eq. \((S=1.3.3)\)), the inverse \((\Sigma_x)^{-1}\) is also a diagonal matrix with diagonal elements given by the following equation:

\[
\{\{\Sigma_x\}^{-1}\}_{j,j} = \frac{1}{\{\Sigma_x\}_{j,j}} = \{\{\Sigma_x\}_{j,j}\}^{-1}. \quad (A3-4)
\]

Then, for an arbitrary \( i \), the following equation is derived from Eqs. \((A3-2)\) and \((A3-4)\):

\[
\{\Sigma_u\}_{i,i} = \sum_{k=1}^{w} C_{i,k} \cdot \left( \sum_{t=1}^{w} \{\{\Sigma_x\}^{-1}\}_{k,t} \cdot C_{t,i}^T \right) = \sum_{k=1}^{w} C_{i,k} \cdot \{\{\Sigma_x\}^{-1}\}_{k,k} \cdot C_{i,k}^T = \sum_{k=1}^{w} C_{i,k}^2 \cdot \{\{\Sigma_x\}_{k,k}\}^{-1}. \quad (A3-5)
\]

Thus, the sum of all elements in the arbitrary \( i \)-th row of the \( CM \) matrix can be calculated using Eqs. \((A3-3)\), \((A3-4)\), and \((A3-5)\) as follows:

\[
\sum_{j=1}^{w} \{CM\}_{i,j} = \sum_{j=1}^{w} \frac{C^2_{i,j}}{\{\Sigma_u\}_{i,i} \cdot \{\Sigma_x\}_{j,j}} = \sum_{j=1}^{w} \frac{C^2_{i,j} \cdot \{\{\Sigma_x\}_{j,j}\}^{-1}}{\sum_{k=1}^{w} C^2_{i,k} \cdot \{\{\Sigma_x\}_{k,k}\}^{-1}} = 1. \quad (A3-6)
\]
References cited in Supplementary File 1

S-1. Wiechert W: \textit{\textsuperscript{13}C metabolic flux analysis}. \textit{Metab Eng} 2001, 3:195-206.

S-2. Varma A, Palsson BO: \textit{Metabolic flux balancing: basic concepts, scientific and practical use}. \textit{Bio/Technology} 1994, 12:994-998.

S-3. Yang TH: \textit{\textsuperscript{13}C-based metabolic flux analysis: fundamentals and practice}. \textit{Methods Mol Biol} 2013, 985:297-334.

S-4. Antoniewicz MR, Kelleher JK, Stephanopoulos G: \textit{Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements}. \textit{Metab Eng} 2006, 8:324-337.

S-5. Foster LV: \textit{Rank and null space calculations using matrix decomposition without column interchanges}. \textit{Linear Algebra Appl} 1986, 74:47-71.

S-6. Quek LE, Wittmann C, Nielsen LK, Krümer JO: \textit{OpenFLUX: efficient modelling software for \textsuperscript{13}C-based metabolic flux analysis}. \textit{Microb Cell Fact} 2009, 8:25.

S-7. Wiechert W, Möllney M, Petersen S, de Graaf AA: \textit{A universal framework for \textsuperscript{13}C metabolic flux analysis}. \textit{Metab Eng} 2001, 3:265-283.

S-8. Yang TH, Frick O, Heinzle E: \textit{Hybrid optimization for \textsuperscript{13}C metabolic flux analysis using systems parametrized by compactification}. \textit{BMC Systems Biol} 2008, 2:29.

S-9. Wiechert W, de Graaf AA: \textit{Bidirectional reaction steps in metabolic networks: I. Modeling and stimulation of carbon isotope labeling experiments}. \textit{Biotechnol Bioeng} 1997, 55:101-117.

S-10. Möllney M, Wiechert W, Kownatzki D, de Graaf AA: \textit{Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments}. \textit{Biotechnol Bioeng} 1999, 66:86-103.

S-11. Wang NS, Stephanopoulos G: \textit{Application of macroscopic balances to the identification of gross measurement errors}. \textit{Biotechnol Bioengin} 1983, 25:2177-2208.

S-12. Yang TH, Bolten CJ, Coppi MV, Sun J, Heinzle E: \textit{Numerical bias estimation for mass spectrometric mass isotopomer analysis}. \textit{Anal Biochem} 2009, 388:192-203.
S-13. Antoniewicz MR, Kelleher JK, Stephanopoulos G: Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab Eng* 2007, 9:68-86.

S-14. Schmidt K, Carlsen M, Nielsen J, Villadsen J: Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices. *Biotechnol Bioeng* 1997, 55:931-840.

S-15. Byrd RH, Hribar ME, Nocedal J: An interior point algorithm for large scale nonlinear programming. *SIAM J Optim* 1999, 9: 877-900.

S-16. Wittmann C, Heinzle E: Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing corynebacteria. *Appl Environ Microbiol* 2002, 68:5843-5859.

S-17. Wittmann C, Kim HM, Heinzle E: Metabolic network analysis of lysine producing *Corynebacterium glutamicum* at a miniaturized scale. *Biotechnol Bioeng* 2004, 87:1-6.

S-18. Leighty RW, Antoniewicz MR: COMPLETE-MFA: Complementary parallel labeling experiments technique for metabolic flux analysis. *Metab Eng* 2013, 20:49-55.

S-19. Bryk SW, Raudenbush AS: Hierarchical linear models: applications and data analysis methods. 2nd edition. Thousand Oaks, CA [u.a.]: Sage Publications; 2002.

S-20. Wiechert W, Siefke C, de Graaf AA, Mark A: Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis. *Biotechnol Bioeng* 1997, 55:118-135.

S-21. Nocedal J, Wright SJ: *Numerical optimization*. 2nd edition. Berlin, New York: Springer-Verlag; 2006.

S-22. Suthers PF, Burgard AP, Dasika MS, Nowroozi F, van Dien S, Keasling JD, Maranas CD: Metabolic flux elucidation for large-scale models using $^{13}$C labeled isotopes. *Metab Eng* 2007, 9:387-405.

S-23. Antoniewicz MR: Using multiple tracers for $^{13}$C metabolic flux analysis. *Methods Mol Biol* 2013, 985:353-365.

S-24. Khoshgoftaar TM, Allen EB: Controlling overfitting in classification-tree models of software quality. *Empirical Software Engineering* 2001, 6:59-79.
S-25. Hartmann WM, Hartwig RE: **Computing the Moore-Penrose inverse for the covariance matrix in constrained nonlinear estimation.** *SIAM J Optim* 1996, **6**:727-747.

S-26. Pázman A: *Foundations of optimum experimental design.* Dordrecht, The Netherlands: Kluwer Academic Publishing; 1986.

S-27. He L, Xiao Y, Gebreselassie N, Zhang F, Antoniewicz MR, Tang YJ, Peng L: **Central metabolic responses to the overproduction of fatty acids in *Escherichia coli* based on $^{13}$C-metabolic flux analysis.** *Biotechnol Bioeng* 2013, **111**:575-85.