Mathematical Biosciences 319 (2020) 108291

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Original research article

Experimental design for parameter estimation in steady-state linear models of metabolic networks

Håvard G. Frøysa*, Hans J. Skaug, Guttorm Alendal

Department of Mathematics, University of Bergen, Bergen, Norway

ARTICLE INFO

Keywords:
Systems biology
Metabolic networks
Parameter identifiability
Experimental design
Fisher information
D-optimality

ABSTRACT

Metabolic networks are typically large, containing many metabolites and reactions. Dynamical models that aim to simulate such networks will consist of a large number of ordinary differential equations, with many kinetic parameters that must be estimated from experimental data. We assume these data to be metabolomics measurements made under steady-state conditions for different input fluxes. Assuming linear kinetics, analytical criteria for parameter identifiability are provided. For normally distributed error terms, we also calculate the Fisher information matrix analytically to be used in the D-optimality criterion. A test network illustrates the developed tool chain for finding an optimal experimental design. The first stage is to verify global or pointwise parameter identifiability, the second stage to find optimal input fluxes, and finally remove redundant measurements.

1. Introduction

Dynamical models of metabolic networks typically consist of a large number of metabolites and reactions (fluxes). Each of the reactions has a reaction rate that is modelled by a rate law parametrized by one or more kinetic parameters. For some reactions and networks, information about these rates can be found in the literature and databases like SABIO-RK[1] and BRENDA[2]. However, this will not generally be the case. To calibrate and validate the simulation models we then have to rely on data from traditional in-vivo, in-vitro and in-situ experiments that, preferably, are designed for this purpose [1,2]. Before performing such experiments, however, the identifiability of the dynamical model should be examined. If a parameter combination is structurally non-identifiable, different parameter values will give the same data such that it is impossible to determine the parameter values even for observations without measurement errors [3].

There is a rich literature on the identifiability of dynamical models for reaction networks. Most papers consider a general dynamical observation function and perform numerical analyses of the identifiability [4–9]. One powerful method is to use the profile likelihood proposed by Raue et al. [10], or the more recent penalized optimization from Kreutz [11]. There are also several studies that consider the identifiability of specific kinetic models, where we mention some of them here. Zou et al. [12] uses the Laplace transform to study the identifiability of a model with dynamic PET data. Gabor et al. [13] use a collinearity index to find the largest possible group of uncorrelated parameters and test the method on several large-scale kinetic models. Casas-Orozco et al. [14] use singular value decomposition to find subsets of identifiable parameters for heterogeneously-catalyzed reactions. Finally, Berthoumieux et al. [15] consider the identifiability for parameters in linlog models of steady-state metabolic networks.

We assume the observations to be steady-state concentrations, and do not require dynamic measurements in contrast to the studies mentioned above. In addition, we assume a linear kinetics structure that makes it possible to find the steady-state concentrations analytically and to view the model as a compartmental model [16]. This enables us to study the identifiability of the parameters by purely analytical methods without the use of any numerical or statistical methods. Our novel criterion for global identifiability of the parameters in Thm. 4 only uses the network structure, and is based on ideas from stoichiometric metabolic flux analysis [17].

Our second main novelty is the analytical calculation of the Fisher information matrix for given parameter values under the assumptions of normally distributed error terms and independence of observations. This enables us to conduct experimental design to find optimal input fluxes and minimal observation sets for an assumed network structure. Having an analytical expression for the Fisher information matrix avoids the numerical problems considered in e.g. Eisenberg et al. [18].

* Corresponding author.
E-mail address: havard.froysa@uib.no (H.G. Frøysa).
[1] http://sabio.villa-bosch.de
[2] http://www.brenda-enzymes.org/

https://doi.org/10.1016/j.mbs.2019.108291
Received 22 March 2019; Received in revised form 20 November 2019; Accepted 20 November 2019
Available online 28 November 2019

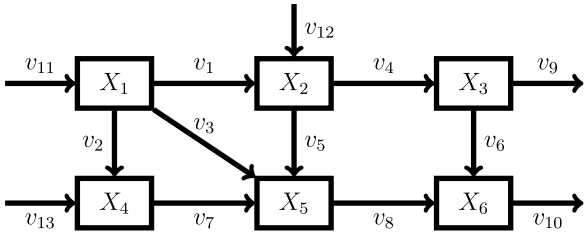

Fig. 1. Example network. The nodes are metabolites with names X_k and concentrations x_k, and the directed edges are irreversible reactions with reaction rates (fluxes) v_j.

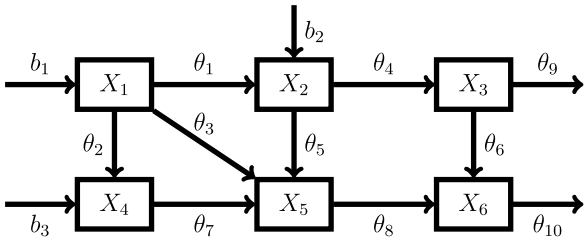

Fig. 2. The network of Fig. 1 with new labels. The nodes (boxes) are metabolites, and the directed edges are irreversible reactions with some flux. The inputs b_j's are known constant reaction rates, while the θ_j's are the first order kinetic parameters to be estimated.

In Sec. 2 we set up the mathematical framework for a reaction network. Identifiability is defined in Sec. 3.1, pointwise identifiability studied in Sec. 3.2 and existence of designs giving identifiability considered in Sec. 3.3. Modified stoichiometric matrices are presented in Sec. 3.4 to prepare for the study of global identifiability in Sec. 3.5. The maximum likelihood problem is formulated in Sec. 3.6, and the Fisher information matrix calculated in Sec. 3.7. Optimal designs are found in Sec. 4.1, and redundant measurements removed in Sec. 4.2. An example is studied in Sec. 5, while Sec. 6 is the discussion.

2. Mathematical modeling of reaction networks

2.1. Reaction network

We start by developing the mathematical framework to be used for a reaction network, where the first step is the stoichiometry of the reactions.

A reaction network consists of n chemical species (metabolites) with names $X_1, \ldots, X_n$ and $\mathbf{x} = [x_1, \ldots, x_n]^{\mathrm{T}} \in \mathbb{R}^n_{\geq 0}$ as their concentrations. The metabolites interact through r irreversible reactions, where pairs may form reversible reactions. We assume that all reactions are of the form

$$\alpha X_s \rightarrow \beta X_p \,, \tag{1}$$

where X_s is the substrate and X_p the product. The stoichiometric constants α and β are non-negative integers assumed to be 0 or 1, which makes the network a compartmental model [16]. The reaction (flux) of Eq. (1) is internal if $[\alpha \;\; \beta] = [1 \;\; 1]$, entering if $[\alpha \;\; \beta] = [0 \;\; 1]$ and leaving if $[\alpha \;\; \beta] = [1 \;\; 0]$.

A network can be written in a notation based on Conradi et al. [19] as

$$\alpha_j X_{i_j^s} \rightarrow \beta_j X_{i_j^p} \qquad j = 1, \ldots, r \quad , \tag{2}$$

where i_j^s is the substrate index of reaction j, and i_j^p is the product index. One of the indices is undefined for entering and leaving fluxes, but never needed.

The network in Fig. 1 will be used as an example throughout the manuscript. It has $n = 6$ metabolites and $r = 13$ reactions listed in Tab. 1. In addition, Appendix C presents a sub-network of Fig. 1 as a simpler and more accessible example of the type of models and results contained in this paper.

A network on the form of Eq. (2) can be represented by a stoichiometric matrix $\mathbf{S} = \{S_{ij}\} \in \mathbb{Z}^{n \times r}$. Entry S_{ij} gives the production of metabolite X_i in reaction j [19]. Column j of $\mathbf{S}$ will have the value $-\alpha_j$ in row i_j^s and β_j in row i_j^p, while the remaining entries are zero. The example in Tab. 1 gives

$$\mathbf{S} = \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{3}$$

2.2. Dynamical model

The next step in the development of the mathematical model for the reaction network is the dynamical equations including the so-called kinetics.

Let $\mathbf{v} = [v_1, \ldots, v_r]^{\mathrm{T}} \in \mathbb{R}^r_{\geq 0}$ be the fluxes (reaction rates) of the reactions. The dynamical behaviour of the concentrations $\mathbf{x}$ is given by the system

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} \tag{4}$$

of ordinary differential equations where the fluxes $\mathbf{v}$ are functions of $\mathbf{x}$, and $\mathbf{S}$ is the constant stoichiometric matrix [20, ch. 9]. For consistency, $\mathbf{x}$ and $\mathbf{v}$ should remain non-negative given non-negative initial conditions. Note that non-negativity of $\mathbf{v}$ is equivalent to irreversibility of all the reactions.

To completely define the dynamical model of Eq. (4) we need to specify $\mathbf{v}$ as a function of $\mathbf{x}$, i.e. the kinetics. We assume that the entering fluxes are constant (zero order kinetics) and known design parameters. Constant fluxes occur when a reaction is saturated, but is also a natural assumption when a flux comes from the outside. The remaining fluxes are assumed to be proportional to the concentration of their substrate (first order kinetics). The proportionality constants are kinetic parameters subject to estimation.

We now separate the fluxes and renumber them if necessary such that

$$\mathbf{S} = [\mathbf{S}_1 \;\; \mathbf{S}_0] \text{ and } \mathbf{v}^{\mathrm{T}} = [\mathbf{v}_1^{\mathrm{T}} \;\; \mathbf{v}_0^{\mathrm{T}}] \,, \tag{5}$$

where 0 refers to the zero order (constant) entering fluxes, and 1 refers to the first order internal and leaving fluxes.

Assuming that there are r_0 zero order and r_1 first order fluxes, where $r_0 + r_1 = r$, we have $\mathbf{S}_0 \in \mathbb{N}_0^{n \times r_0}$, $\mathbf{S}_1 \in \mathbb{Z}^{n \times r_1}$, $\mathbf{v}_0 \in \mathbb{R}^{r_0}_{\geq 0}$ and $\mathbf{v}_1 \in \mathbb{R}^{r_1}_{\geq 0}$. The reactions in Tab. 1 are numbered this way with $r_0 = 3$ and $r_1 = 10$ such that

$$\mathbf{S}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_1$$

$$= \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \end{bmatrix}. \tag{6}$$

Let $\mathbf{b} = [b_1 \;\; \ldots \;\; b_{r_0}]^{\mathrm{T}} \in \mathbb{R}^{r_0}_{\geq 0}$ be the r_0 input fluxes such that $\mathbf{v}_0 = \mathbf{b}$. Further, let $\boldsymbol{\theta} = [\theta_1 \;\; \ldots \;\; \theta_{r_1}]^{\mathrm{T}} \in \mathbb{R}^{r_1}_{>0}$ be the r_1 kinetic parameters such that

$$\mathbf{v}_1 = \left[\theta_1 x_{i_1^s} \quad \ldots \quad \theta_{r_1} x_{i_{r_1}^s}\right]^{\mathrm{T}}, \tag{7}$$

where we see that $\mathbf{v}_1$ is a bilinear function of $\boldsymbol{\theta}$ and $\mathbf{x}$. Note that all θ_j

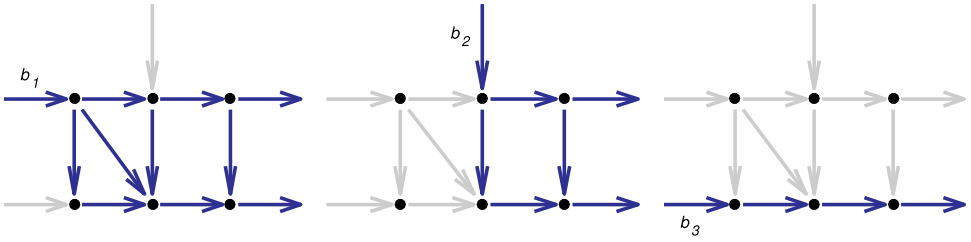

Fig. 3. Active reactions for the example and the standard input vectors $\mathbf{e}_j$ associated with the inputs b_j. Dark blue arrows are active, and light gray arrows inactive. See Fig. 1 for labelling of the network with fluxes v_j and Fig. 2 for labelling with parameters θ_j.

Table 1
Reactions for the example in Fig. 1 on the form of Eq. (2).

$j=1$	X_1	$\rightarrow$	X_2	$i_1^s=1$	$i_1^p=2$
$j=2$	X_1	$\rightarrow$	X_4	$i_2^s=1$	$i_2^p=4$
$j=3$	X_1	$\rightarrow$	X_5	$i_3^s=1$	$i_3^p=5$
$j=4$	X_2	$\rightarrow$	X_3	$i_4^s=2$	$i_4^p=3$
$j=5$	X_2	$\rightarrow$	X_5	$i_5^s=2$	$i_5^p=5$
$j=6$	X_3	$\rightarrow$	X_6	$i_6^s=3$	$i_6^p=6$
$j=7$	X_4	$\rightarrow$	X_5	$i_7^s=4$	$i_7^p=5$
$j=8$	X_5	$\rightarrow$	X_6	$i_8^s=5$	$i_8^p=6$
$j=9$	X_3	$\rightarrow$		$i_9^s=3$	
$j=10$	X_6	$\rightarrow$		$i_{10}^s=6$	
$j=11$		$\rightarrow$	X_1		$i_{11}^p=1$
$j=12$		$\rightarrow$	X_2		$i_{12}^p=2$
$j=13$		$\rightarrow$	X_4		$i_{13}^p=4$

and b_j must be non-negative to ensure irreversibility of all the reactions. However, we require all θ_j to be strictly positive such that an internal or leaving flux is non-zero if its substrate concentration is non-zero. For the example we get

$$\mathbf{v}_0 = [b_1 \quad b_2 \quad b_3]^{\mathrm{T}} \text{ and}$$
$$\mathbf{v}_1 = [\theta_1 x_1 \quad \theta_2 x_1 \quad \theta_3 x_1 \quad \theta_4 x_2 \quad \theta_5 x_2 \quad \theta_6 x_3 \quad \theta_7 x_4 \quad \theta_8 x_5 \quad \theta_9 x_3 \quad \theta_{10} x_6]^{\mathrm{T}}, \tag{8}$$

which we use in Fig. 2 to redraw Fig. 1 with the b_j's and θ_j's as labels.

The bilinear structure of $\mathbf{v}_1$ makes it possible to write the fluxes as

$$\mathbf{v}_0 = \mathbf{b} \quad \text{and} \quad \mathbf{v}_1 = \mathbf{K}_\theta \mathbf{x}, \tag{9}$$

where the matrix $\mathbf{K}_\theta = \{K_{ji}\} \in \mathbb{R}_{\geq 0}^{r_1 \times n}$ is given by the parameters in $\boldsymbol{\theta}$ as $K_{ji} = \delta_{i,i_j^s}\theta_j$, and $\delta_{i,j}$ is the Kronecker-delta

$$\delta_{i,j} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}. \tag{10}$$

Each row j of $\mathbf{K}_{\boldsymbol{\theta}}$ is associated with a reaction. The only non-zero entry of this row is θ_j in column i_j^s, the substrate index. For the example we get

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{and} \quad \mathbf{K}_\theta = \begin{bmatrix} \theta_1 & 0 & 0 & 0 & 0 & 0 \\ \theta_2 & 0 & 0 & 0 & 0 & 0 \\ \theta_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_4 & 0 & 0 & 0 & 0 \\ 0 & \theta_5 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_6 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_7 & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_8 & 0 \\ 0 & 0 & \theta_9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{10} \end{bmatrix}. \tag{11}$$

Combining the flux vectors into one vector according to Eq. (5) gives

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_\theta \\ \boldsymbol{0} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \boldsymbol{0} \\ \mathbf{b} \end{bmatrix}, \tag{12}$$

such that Eq. (4) becomes the inhomogeneous linear ODE system

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} = [\mathbf{S}_1 \quad \mathbf{S}_0]\left(\begin{bmatrix} \mathbf{K}_\theta \\ \boldsymbol{0} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \boldsymbol{0} \\ \mathbf{b} \end{bmatrix}\right) = \mathbf{D}_\theta \mathbf{x} + \mathbf{c}, \tag{13}$$

with $\mathbf{D}_\theta = \mathbf{S}_1\mathbf{K}_\theta \in \mathbb{R}^{n \times n}$ and $\mathbf{c} = \mathbf{S}_0\mathbf{b} \in \mathbb{R}_{\geq 0}^n$. The example has

$$\mathbf{D}_\theta = \begin{bmatrix} -\theta_1 - \theta_2 - \theta_3 & 0 & 0 & 0 & 0 & 0 \\ \theta_1 & -\theta_4 - \theta_5 & 0 & 0 & 0 & 0 \\ 0 & \theta_4 & -\theta_6 - \theta_9 & 0 & 0 & 0 \\ \theta_2 & 0 & 0 & -\theta_7 & 0 & 0 \\ \theta_3 & \theta_5 & 0 & \theta_7 & -\theta_8 & 0 \\ 0 & 0 & \theta_6 & 0 & \theta_8 & -\theta_{10} \end{bmatrix}$$

$$\text{and} \quad \mathbf{c} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ b_3 \\ 0 \\ 0 \end{bmatrix}. \tag{14}$$

The matrix $\mathbf{D}_{\boldsymbol{\theta}}$ is a compartmental matrix [16], where the metabolites are the compartments. It has non-negative off-diagonal entries and non-positive diagonal entries. This ensures non-negativity of $\mathbf{x}$ for all t in the homogeneous system corresponding to Eq. (13) given non-negative initial values [16, p. 48]. Since $\mathbf{c}$ has only non-negative entries, non-negativity of $\mathbf{x}$ for all t given non-negative initial values is also satisfied for Eq. (13).

All eigenvalues λ of $\mathbf{D}_{\boldsymbol{\theta}}$ have $\mathrm{Re}(\lambda) \leq 0$ by the Gerschgorin circle theorem applied to the columns of $\mathbf{D}_{\boldsymbol{\theta}}$ since it is a compartmental matrix [16, p. 55]. Also, zero is not an eigenvalue of $\mathbf{D}_{\boldsymbol{\theta}}$ if all the metabolites (compartments) have an exit to the outside environment [16, p. 56]. This implies that $\mathrm{Re}(\lambda) < 0$ for all λ making $\mathbf{D}_{\boldsymbol{\theta}}$ invertible and the dynamical system of Eq. (13) asymptotically stable for all initial conditions of $\mathbf{x}$. Finally, it also implies that the inverse $\mathbf{D}_\theta^{-1}$ has only non-positive entries [16, p. 69].

2.3. Steady-state conditions

We now have a dynamical model for a reaction network, and will later introduce observations of the concentrations. It is assumed that they are made in steady-state conditions [21, p. 114], imposing the flux balance

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} = \mathbf{0}. \tag{15}$$

Hence, the flux vector $\mathbf{v}$ should be in the nullspace of $\mathbf{S}$ where the matrix usually has more columns (reactions) than rows (metabolites). This guarantees a non-trivial nullspace and gives infinitely many possible steady-states.

The steady-state condition of Eq. (15) applied to Eq. (13) gives

$$\mathbf{D}_\theta \mathbf{x} + \mathbf{c} = \boldsymbol{0}, \tag{16}$$

which is a linear equation system in $\mathbf{x}$ for given $\mathbf{c}$ and $\mathbf{D}_\theta$. Since $\mathbf{D}_\theta^{-1}$ exists, there is a unique steady-state $\mathbf{z} \in \mathbb{R}^n_{\geq 0}$ for Eq. (16) given by

$$\mathbf{z} = -\mathbf{D}_\theta^{-1}\mathbf{c}\,. \tag{17}$$

Non-negativity of $\mathbf{z}$ is guaranteed since $\mathbf{D}_\theta^{-1}$ has non-positive and $\mathbf{c}$ non-negative entries. Remember that the dynamical system of Eq. (13) is asymptotically stable for all initial conditions of $\mathbf{x}$, implying $\lim_{t\to\infty}\mathbf{x} = \mathbf{z}$.

If $\mathbf{z}$ and $\mathbf{c}$ are known, the bilinear structure of $\mathbf{v}_1$ seen in Eq. (7) enables us to write the steady-state condition of Eq. (15) as a linear equation system in $\boldsymbol{\theta}$ instead of Eq. (16). To do so, we write the fluxes $\mathbf{v}_1 = \mathbf{X}\theta$ where the diagonal matrix $\mathbf{X} \in \mathbb{R}^{r_1\times r_1}_{\geq 0}$ is given by the concentrations as $\mathbf{X} = \mathrm{diag}(x_{i_1^s},...,x_{i_{r_1}^s})$.

As a linear system in $\boldsymbol{\theta}$, the dynamical system of Eq. (4) becomes

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}_1\mathbf{X}\theta + \mathbf{c} = \mathbf{E}_\mathbf{x}\theta + \mathbf{c} \tag{18}$$

where the entries of the matrix $\mathbf{E}_\mathbf{x} \in \mathbb{R}^{n\times r_1}$ are linear combinations of the concentrations x_k. The steady-state condition of Eq. (15) can now be written

$$\mathbf{E}_\mathbf{z}^\theta\theta = -\mathbf{c} \tag{19}$$

where $\mathbf{E}_\mathbf{z}^\theta$ is $\mathbf{E}_\mathbf{x}$ with $\mathbf{x} = \mathbf{z}$, which is a linear equation system in $\boldsymbol{\theta}$ that we will make use of later. Note the superscript $\boldsymbol{\theta}$ for $\mathbf{E}_\mathbf{z}^\theta$ since $\mathbf{E}_\mathbf{x}$ depends on $\boldsymbol{\theta}$ when $\mathbf{x} = \mathbf{z}$. For the example, we have the matrices

$$\mathbf{X} = \mathrm{diag}(x_1, x_1, x_1, x_2, x_2, x_3, x_4, x_5, x_3, x_6) \quad \text{and} \tag{20}$$

$$\mathbf{E}_\mathbf{x} = \mathbf{S}_1\mathbf{X} = \begin{bmatrix} -x_1 & -x_1 & -x_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ x_1 & 0 & 0 & -x_2 & -x_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & 0 & -x_3 & 0 & 0 & -x_3 & 0 \\ 0 & x_1 & 0 & 0 & 0 & 0 & -x_4 & 0 & 0 & 0 \\ 0 & 0 & x_1 & 0 & x_2 & 0 & x_4 & -x_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & x_3 & 0 & x_5 & 0 & -x_6 \end{bmatrix}. \tag{21}$$

The matrix $\mathbf{E}_\mathbf{z}^\theta$ will be important in the later sections about identifiability.

3. Parameter estimation

We will now consider the estimation of the kinetic parameters in the vector $\theta = [\theta_1,...,\theta_{r_1}]^\mathrm{T} \in \mathbb{R}^{r_1}_{>0}$. The data are observations of the steady-state concentrations $\mathbf{z}$ for different known input vectors $\mathbf{b}$, subject to measurement error. First, we will look at pointwise and global identifiability of $\boldsymbol{\theta}$ for a given experimental setup without considering measurement errors. Thereafter, we define the likelihood function and compute the Fisher information matrix analytically. Finally, this matrix is used for optimal design of the input vectors $\mathbf{b}$ and to determine which concentrations that should be observed.

3.1. Identifiability

It is assumed that we have a set $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ of m different input flux vectors $\mathbf{b}$ called a design, and that the parameters in $\boldsymbol{\theta}$ are the same under all of these. Let $\mathbf{c}^l = \mathbf{S}_0\mathbf{b}^l$ and $\mathbf{z}^l = [z_1^l...z_n^l]^\mathrm{T} = -\mathbf{D}_\theta^{-1}\mathbf{c}^l$ be the vectors $\mathbf{c}$ and $\mathbf{z}$ for the different $\mathbf{b}^l$.

Identifiability is the property of the mapping $\boldsymbol{\theta} \to \mathbf{z}(\boldsymbol{\theta})$ given by Eq. (17) and a design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ that distinct values of $\boldsymbol{\theta}$ give distinct data sets.

Definition 1. The parameter vector $\boldsymbol{\theta}$ is *pointwise identifiable* at $\theta_1 \in \mathbb{R}^{r_1}_{>0}$ for a given design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ if

$$\theta_1 \neq \theta_2 \quad \Rightarrow \quad \mathbf{z}^l(\theta_1) \neq \mathbf{z}^l(\theta_2) \text{ for some } l \in \{1,...,m\} \tag{22}$$

holds for all $\theta_2 \in \mathbb{R}^{r_1}_{>0}$.

To illustrate pointwise identifiability for the example of Fig. 2 we look at the design $\{\mathbf{b}^1, \mathbf{b}^2\}$ with the $m = 2$ input vectors

$$\mathbf{b}^1 = [1\ \ 0\ \ 0]^\mathrm{T} \quad \text{and} \quad \mathbf{b}^2 = [0\ \ 1\ \ 2]^\mathrm{T}. \tag{23}$$

For the parameter values $\theta_1 = [2\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1]^\mathrm{T}$, the equalities

$$\mathbf{z}^1(\theta_1) = \mathbf{z}^1(\theta_2) \quad \text{and} \quad \mathbf{z}^2(\theta_1) = \mathbf{z}^2(\theta_2) \tag{24}$$

only hold for $\theta_1 = \theta_2$, giving pointwise identifiability at $\boldsymbol{\theta}_1$. However, for the values $\theta_1 = [1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1]^\mathrm{T}$, the equalities of Eq. (24) hold for all

$$\begin{aligned} &\theta_2 = \theta_1 + \gamma_1\mathbf{a}_1 + \gamma_2\mathbf{a}_2 \quad \text{where} \\ &\mathbf{a}_1^\mathrm{T} = [0\ \ 0\ \ 0\ \ -5\ \ 5\ \ -10\ \ 0\ \ 1\ \ 0\ \ 0]\,, \\ &\mathbf{a}_2^\mathrm{T} = [0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 11\ \ 0\ \ 0\ \ -11\ \ 1] \end{aligned} \tag{25}$$

and $\gamma_1, \gamma_2 \in \mathbb{R}$. This can be found by the use of Eq. (28), and means that we don't have pointwise identifiability at this $\boldsymbol{\theta}_1$. Two more similar examples are given in Appendix B.

Altogether, this illustrates that for a given design there may exist values of $\boldsymbol{\theta}$ with pointwise identifiability and other values without. For this reason, we also need to look at global identifiability.

Definition 2. The parameter vector $\boldsymbol{\theta}$ is *globally identifiable* for a given design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ if Def. 1 for pointwise identifiability holds for all $\theta_1 \in \mathbb{R}^{r_1}_{>0}$, which is equivalent to the mapping $\theta \to (\mathbf{z}^1(\theta),...,\mathbf{z}^m(\theta))$ being injective.

Note that the identifiability depends on the design $\{\mathbf{b}^1,...,\mathbf{b}^m\}$, which is thought of as fixed in this section. Thus, the definition of global identifiability is global in $\boldsymbol{\theta}$, but in a sense pointwise in the design. However, we will later find optimal designs and examine how the identifiability depends on the design. It is particularly interesting whether or not there exists a design that provides pointwise or global identifiability.

3.2. Pointwise identifiability criterion

We now give a criterion for pointwise identifiability of $\boldsymbol{\theta}$ for a design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ by utilizing the steady-state formulation of Eq. (19).

Let the matrix $\mathbf{E}^\theta \in \mathbb{R}^{mn\times r_1}$ and the vector $\mathbf{C} \in \mathbb{R}^{mn}_{\geq 0}$ be

$$\mathbf{E}^\theta = \begin{bmatrix} \mathbf{E}_{\mathbf{z}^1}^\theta \\ \vdots \\ \mathbf{E}_{\mathbf{z}^m}^\theta \end{bmatrix} \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} \mathbf{c}^1 \\ \vdots \\ \mathbf{c}^m \end{bmatrix} \tag{26}$$

where $\mathbf{E}_{\mathbf{z}^l}^\theta$ is $\mathbf{E}_\mathbf{z}^\theta$ from Eq. (19) with $\mathbf{z} = \mathbf{z}^l$, and $\mathbf{c}^l = \mathbf{S}_0\mathbf{b}^l$ as before.

Theorem 1. *The parameter vector $\theta \in \mathbb{R}^{r_1}_{>0}$ is pointwise identifiable at $\theta = \theta_1$ for a given design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ if and only if* $\mathrm{Rank}(\mathbf{E}^{\theta_1}) = r_1$.

Proof. The true parameter vector $\boldsymbol{\theta}_1$ satisfies the m equilibrium conditions

$$\mathbf{E}_{\mathbf{z}^l}^{\theta_1}\theta_1 = -\mathbf{c}^l \tag{27}$$

of Eq. (19) for all l by construction. This gives the linear system

$$\mathbf{E}^{\theta_1}\theta = -\mathbf{C} \tag{28}$$

of equations in $\boldsymbol{\theta}$, which by construction has the true parameter values $\boldsymbol{\theta}_1$ as a solution. To have pointwise identifiability, $\mathrm{Null}(\mathbf{E}^{\theta_1})$ must be trivial such that the true parameter vector $\boldsymbol{\theta}_1$ is the unique solution of Eq. (28). This is equivalent to $\mathbf{E}^{\theta_1}$ having full column rank r_1. □

The number of equations in Eq. (28) could be much larger than the number of unknowns, but the system will always have a solution by construction. If the nullspace is non-trivial, some parameters may still be uniquely determined by Eq. (28). Also be aware that the identifiability is a property of the design $\{\mathbf{b}^1,...,\mathbf{b}^m\}$, not of the individual vectors $\mathbf{b}^l$.

3.3. Existence of identifiability design

An important question is whether or not there exists a design $\{\mathbf{b}^1, ..., \mathbf{b}^m\}$ that provides identifiability. We will now show that this can be checked by considering the design $\{\mathbf{e}_1, ..., \mathbf{e}_{r_0}\}$ of standard basis vectors in $\mathbb{R}^{r_0}$.

Theorem 2. *There exists a design giving pointwise identifiability at $\theta = \theta_1$ if and only if the design $\{\mathbf{e}_1, ..., \mathbf{e}_{r_0}\}$ gives pointwise identifiability at $\theta = \theta_1$.*

Proof. We only need to prove that pointwise identifiability for an arbitrary design implies pointwise identifiability for the standard design.

The steady-state concentrations $\mathbf{z}$ for an input vector $\mathbf{b}$ are given by the linear relation $\mathbf{z} = -\mathbf{D}_\theta^{-1}\mathbf{c} = -\mathbf{D}_\theta^{-1}\mathbf{S}_0\mathbf{b}$ of Eq. (17). We can write $\mathbf{b}$ as

$$\mathbf{b} = [b_1 \ \ ... \ \ b_{r_0}]^{\mathrm{T}} = \sum_{j=1}^{r_0} b_j \mathbf{e}_j \tag{29}$$

where $\mathbf{e}_j$ is the j'th standard basis vector in $\mathbb{R}^{r_0}$ as before. For this reason, any steady-state concentration vector $\mathbf{z}$ is a linear combination

$$\mathbf{z} = -\sum_{j=1}^{r_0} b_j \mathbf{D}_\theta^{-1}\mathbf{S}_0\mathbf{e}_j \tag{30}$$

given by the entries b_j in $\mathbf{b}$. Let $\mathbf{E}_j^\theta$ be the matrix $\mathbf{E}_\mathbf{z}^\theta$ where $\mathbf{z} = -\mathbf{D}_\theta^{-1}\mathbf{S}_0\mathbf{e}_j$, and let $\mathbf{E}_S^\theta$ be $\mathbf{E}^\theta$ of Eq. (26) for the standard design $\{\mathbf{e}_1, ..., \mathbf{e}_{r_0}\}$. Since $\mathbf{E}_\mathbf{z}^\theta$ is linear in $\mathbf{z}$ and $\mathbf{z}$ is linear in $\mathbf{b}$, also $\mathbf{E}_\mathbf{z}^\theta$ will be linear in $\mathbf{b}$ such that

$$\mathbf{E}_\mathbf{z}^\theta = \sum_{j=1}^{r_0} b_j \mathbf{E}_j^\theta \tag{31}$$

for any steady-state $\mathbf{z}$. This implies that the set $\{\mathbf{E}_1^\theta, ..., \mathbf{E}_{r_0}^\theta\}$ spans the space of matrices $\mathbf{E}_\mathbf{z}^\theta$ for steady-states $\mathbf{z}$. A consequence of this is that

$$\mathrm{Row}(\mathbf{E}^\theta) \subseteq \mathrm{Row}(\mathbf{E}_S^\theta) \tag{32}$$

for the matrix $\mathbf{E}^\theta$ of any given design. Then, for the nullspaces we get

$$\mathrm{Null}(\mathbf{E}_S^\theta) \subseteq \mathrm{Null}(\mathbf{E}^\theta), \tag{33}$$

where the space on the right side of Eq. (33) must trivial for $\theta = \theta_1$ to have pointwise identifiability at $\boldsymbol{\theta}_1$ for a design $\{\mathbf{b}^1, ..., \mathbf{b}^m\}$. It then follows that the space on the left side of Eq. (33) must be trivial such that the standard design gives pointwise identifiability. □

If there exists a design that provides pointwise identifiability at $\theta = \theta_1$, we have $\mathrm{Rank}(\mathbf{E}_S^{\theta_1}) = n_1$ by Thm. 1 and 2. Note, however, that the pointwise identifiability can also be examined by other approaches than Thm. 1, including the Fisher information matrix to be introduced later.

3.4. Active reactions and modified stoichiometric matrices

Next, we want to find a criteria for global identifiability of $\boldsymbol{\theta}$. To prepare for this, we classify reactions as active or non-active and define modified stoichiometric matrices to be used in the global identifiability criterion.

A reaction with a non-zero steady-state flux $v_j > 0$ under input vector $\mathbf{b}$ is said to be an active reaction for the input vector. Likewise, a metabolite with a non-zero steady-state concentration $z_k > 0$ under input vector $\mathbf{b}$ is said to be an active metabolite for the input vector. Note that a reaction j is active if and only if the substrate is an active metabolite, i.e. if $z_{i_j^s} > 0$.

Remember that the steady-state concentrations $\mathbf{z}$ can be calculated from Eq. (17) for a given $\mathbf{b}$ and $\boldsymbol{\theta}$. However, the set of active metabolites and reactions for a given input vector $\mathbf{b}$ will be the same for all $\boldsymbol{\theta}$ since the parameters in $\boldsymbol{\theta}$ are assumed to be strictly positive. It will also be the same for all scalings of the input vector $\mathbf{b}$ due to the linear structure of Eq. (17).

Table 2
The active reactions for the standard input vectors $\mathbf{e}_j$ of the example in Fig. 2, marked with a cross if the reaction is active. The same information is visualized in Fig. 3.

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}
$\mathbf{e}_1$	×	×	×	×	×	×	×	×	×	×
$\mathbf{e}_2$				×	×	×		×	×	×
$\mathbf{e}_3$							×	×		×

Table 3
Optimal designs of m input vectors for the example of Fig. 2 with the parameter values of Eq. (56). Each design is the solution of Eq. (50) with the constraint of Eq. (51). The last column is whether or not the design gives global identifiability checked by Thm. 4.

m	$\mathbf{b}_*^1, ..., \mathbf{b}_*^m$	$\mathrm{Tr}(\mathbf{I}_{\bar{\theta}}^{-1})\cdot\sigma^{-2}$	$\det(\mathbf{I}_\theta)\cdot\sigma^{20}$	Global
2	$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0.838\\0.162\end{bmatrix}$	$5.97 \cdot 10^2$	$7.07{\cdot}10^{-13}$	No
3	$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$	$2.03 \cdot 10^2$	$3.35{\cdot}10^{-9}$	Yes
4	$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$	$1.63 \cdot 10^2$	$5.00{\cdot}10^{-8}$	Yes
5	$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$	$1.14 \cdot 10^2$	$6.43{\cdot}10^{-7}$	Yes
6	$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$	$1.02 \cdot 10^2$	$3.43{\cdot}10^{-6}$	Yes

Table 4
Minimal observation set Ω of Eq. (40) for the example in Fig. 2 and the $m = 3$ standard design of Tab. 3. Elements (l, k) in Ω are marked with a cross, and correspond to concentrations z_k^l of metabolite X_k under the standard basis input vector $\mathbf{e}_l$. The included elements are the active metabolites with $z_k^l > 0$ as seen from Fig. 3 or Eq. (57).

$l \setminus k$	1	2	3	4	5	6
1	×	×	×	×	×	×
2		×	×		×	×
3				×	×	×

F or small networks like the one in Fig. 2, the active metabolites and reactions can be identified by inspection, as the metabolites and reactions are active if they have a path from a non-zero input leading to them. The active reactions for the example and the standard basis vectors $\mathbf{e}_j$ are given in Tab. 2 and visualised in Fig. 3, while the active metabolites are given in Tab. 4. For an arbitrary input vector $\mathbf{b}$ will the set of active reactions/metabolites be the union of the equivalent sets for each of standard basis vectors $\mathbf{e}_j$ that are associated with a non-zero entry in $\mathbf{b}$.

We now use the information about the active and non-active reactions for an input vector $\mathbf{b}$. By utilizing the splitting of $\mathbf{S}$ and $\mathbf{v}$ from Eq. (5), the steady-state condition of Eq. (15) can be written

$$\mathbf{S}_1\mathbf{v}_1 = -\mathbf{c} \tag{34}$$

where $\mathbf{c} = \mathbf{S}_0\mathbf{b}$, as before. Let $\mathbf{S}_1^l$ be the matrix $\mathbf{S}_1$ where the columns of non-active reactions for $\mathbf{b}^l$ are set to zero. For the example and the standard basis design $\mathbf{b}^l = \mathbf{e}_l$ with $l \in \{1, 2, 3\}$, we get

$$\mathbf{S}_1^1 = \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \end{bmatrix},$$

$$\mathbf{S}_1^2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -1 & -1 & 0 & \mathbf{0} & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 & 0 & -1 & \mathbf{0} & 0 & -1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 1 & 0 & \mathbf{0} & -1 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 1 & \mathbf{0} & 1 & 0 & -1 \end{bmatrix} \text{ and}$$

$$\mathbf{S}_1^3 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -1 & 0 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 & -1 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 1 & \mathbf{0} & -1 \end{bmatrix} \tag{35}$$

where a light gray column is a non-active reaction in accordance with Tab. 2.

Since all the non-active reactions have a zero flux we can replace the steady-state condition of Eq. (34) for $\mathbf{v}_1 = \mathbf{v}_1^l$ under $\mathbf{b}^l$ with

$$\mathbf{S}_1^l\mathbf{v}_1^l = -\mathbf{c}^l \tag{36}$$

where the active fluxes are balanced using only the active reactions.

Note that there are no restrictions on the non-active fluxes in Eq. (36), and that any value of these will satisfy Eq. (36) even though we know that they are zero. However, the restrictions on the active fluxes are stronger than in Eq. (34), making Eq. (36) important in the global identifiability analysis.

3.5. Global identifiability criterion

We now explore the global identifiability of $\boldsymbol{\theta}$. As for pointwise identifiability, we can restrict ourselves to the standard basis design to determine if global identifiability is possible.

Theorem 3. *There exists a design that gives global identifiability of $\boldsymbol{\theta}$ if and only if the design $\{\mathbf{e}_1, \ldots, \mathbf{e}_{r_0}\}$ gives global identifiability of $\boldsymbol{\theta}$.*

Proof. Global identifiability is equivalent to pointwise identifiability for all values of $\boldsymbol{\theta}$. The theorem follows from Thm. 2 for pointwise identifiability. □

Next, we will see that the global identifiability of $\boldsymbol{\theta}$ can be determined by the matrices $\mathbf{S}_1^l$ of Sec. 3.4. Let the matrix $\mathbf{S}_P \in \mathbb{Z}^{mn \times r_1}$ be given by

$$\mathbf{S}_P = \begin{bmatrix} \mathbf{S}_1^1 \\ \vdots \\ \mathbf{S}_1^m \end{bmatrix} \tag{37}$$

where $\mathbf{S}_1^l$ is the modified stoichiometric matrix for the input vector $\mathbf{b}^l$.

Based on this matrix we can provide a criterion for global identifiability by using ideas from Metabolic Flux Analysis (MFA) [17].

Theorem 4. *The parameter vector $\theta \in \mathbb{R}_{>0}^{r_1}$ is globally identifiable for the design $\{\mathbf{b}^1, \ldots, \mathbf{b}^m\}$ if and only if* $\mathrm{Rank}(\mathbf{S}_P) = r_1$ *where $\mathbf{S}_P$ is from Eq. (37).*

Proof. Let $\mathbf{v}_1^l$ be the first order steady-state fluxes $\mathbf{v}_1$ for the input vector $\mathbf{b}^l$. Each of the vectors satisfies Eq. (36) such that we get the equation system

$$\begin{cases} \mathbf{S}_1^1\mathbf{v}_1^1 = -\mathbf{c}^1 \\ \quad\vdots \\ \mathbf{S}_1^m\mathbf{v}_1^m = -\mathbf{c}^m \end{cases} \tag{38}$$

for the fluxes $\mathbf{v}_1^l = [v_1^l \ldots v_{r_1}^l]^{\mathrm{T}}$ that are unknown when doing inference about $\boldsymbol{\theta}$. However, all these fluxes are given by $v_j^l = \theta_j z_{ij}^l$ where the steady-state concentrations z_k^l are assumed known. This implies that to find θ_j is equivalent to determining v_j^l for a $\mathbf{b}^l$ where the reaction is active with $z_{ij}^l > 0$ as defined in Sec. 3.4. Also remember that $\boldsymbol{\theta}$ is assumed the same under all vectors $\mathbf{b}^l$. If v_j^l is determined for one vector $\mathbf{b}^l$ providing θ_j, the value of v_j^l for all other l can then be computed as $v_j^l = \theta_j z_{ij}^l$. Altogether, identifiability of θ_j is equivalent to v_j^l being determinable for some $l \in \{1, \ldots, m\}$. Thus, v_j^l has to be indeterminable for all $l \in \{1, \ldots, m\}$ if θ_j is unidentifiable. If one or more fluxes v_j^l are indeterminable for all l, there exist a non-trivial $\mathbf{v}_1 \in \mathbb{R}^{r_1}$ such that $\mathbf{S}_1^l\mathbf{v}_1 = \mathbf{0}$ for all $l \in \{1, \ldots, m\}$. These m vector equations for $\mathbf{v}_1$ can be written in the combined form

$$\mathbf{S}_P\mathbf{v}_1 = \mathbf{0} \tag{39}$$

where $\mathbf{S}_P$ is given by Eq. (37). A non-trivial $\mathbf{v}_1$ satisfying Eq. (39) exists if and only if $\mathbf{S}_P$ does not have full column rank r_1. Thus, all parameters θ_j are identifiable if $\mathrm{Rank}(\mathbf{S}_P) = r_1$.

On the other hand, there is no other information available to determine the active fluxes of the various $\mathbf{v}_1^l$ than the steady-state equations of Eq. (38). This implies that the criterion of this theorem must be satisfied to obtain identifiability. Since no specific value of $\boldsymbol{\theta}$ is used, the identifiability is global. □

By combining Thm. 3 and 4, we see that it is sufficient to check the rank of $\mathbf{S}_P$ for the standard design to determine if global identifiability is possible for a given network. For the example, we have $\mathrm{Rank}(\mathbf{S}_P) = r_1 = 10$ for the standard design from the matrices in Eq. (35) such that this is possible.

Note that all $\mathbf{v}_1^l$ satisfy the original steady-state condition $\mathbf{S}_1\mathbf{v}_1^l = -\mathbf{c}^l$ from Eq. (34) in addition to Eq. (36). However, it is the use of the different matrices $\mathbf{S}_1^l$ in Eq. (38) that makes it possible to formulate the simple criterion of Thm. 4. Also note that it is possible to have pointwise identifiability for special designs and values of $\boldsymbol{\theta}$ even though the design does not provide global identifiability, as seen for the design in Eq. (23).

3.6. Maximum likelihood estimation (MLE)

We now formulate a maximum likelihood estimation for $\boldsymbol{\theta}$. Thus far we have assumed that all steady-state concentrations z_k^l are known free of error where $k \in \{1, \ldots, n\}$ is the index of the metabolites and $l \in \{1, \ldots, m\}$ is the index of the input vectors $\mathbf{b}^l$. However, in a real world experiment we will not be able to observe the exact value of a steady-state concentration z_k^l, calling for a statistical model.

For each of the concentrations z_k^l, let y_k^l be the corresponding observed value including measurement error with vector notation $\mathbf{y}^l = [y_1^l \ldots y_n^l]^{\mathrm{T}}$. In addition, it may not be possible or necessary to observe all the concentrations. For this reason, define Ω to be the observation set

$$\Omega = \{(l, k) | \text{ an observation } y_k^l \text{ of } z_k^l \text{ is available}\}. \tag{40}$$

In the previous identifiability discussion it was implicit that $(l, k) \in \Omega$ for all l and k, i.e. that all observations are made. However, we will later explore for which l and k one can have $(l, k) \notin \Omega$ without altering the identifiability.

We assume that the observation y_k^l is a realization of a stochastic variable Y_k^l that is normally distributed, $Y_k^l \sim N(z_k^l, \sigma^2)$ with a common variance σ^2. All the Y_k^l's are also assumed to be mutually independent for all k and l.

Under these assumptions, the marginal probability density of Y_k^l is

$$f_k^l(y_k^l | \boldsymbol{\theta}, \mathbf{b}^l, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(y_k^l - z_k^l)^2\right\} \tag{41}$$

and the joint density f for all the Y_k^l's is the product

$$f(\mathbf{y}^1,...,\mathbf{y}^m|\theta,\mathbf{b}^1,...,\mathbf{b}^m,\sigma^2) = \prod_{(l,k)\in\Omega} f_k^l(y_k^l|\theta,\mathbf{b}^l,\sigma^2). \tag{42}$$

Note that only the observed concentrations are included in Eq. (42) as defined by Ω in Eq. (40). It follows that the log-likelihood function is

$$\ln f = -\frac{1}{2}\sum_{(l,k)\in\Omega}\left[\ln(2\pi\sigma^2) + \frac{(y_k^l - z_k^l)^2}{\sigma^2}\right]. \tag{43}$$

The estimate $\hat{\theta}$ of $\boldsymbol{\theta}$ for a given data set $\{\mathbf{y}^1, ...,\mathbf{y}^m\}$ is then

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}[\ln f(\mathbf{y}^1,...,\mathbf{y}^m|\theta,\mathbf{b}^1,...,\mathbf{b}^m,\sigma^2)] \tag{44}$$

where the design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$ is assumed fixed and given. This will be a non-linear regression problem, see e.g. Bates et al. [22] for an introduction.

In practice, numerical optimization is required to obtain the estimate $\hat{\theta}$ from Eq. (44). If the model is not globally identifiable, the numerical routine can run into problems that are hard to diagnose. Hence, prior to solving Eq. (44) one should check for global identifiability. If this is satisfied, the mapping $\theta \to (\mathbf{z}^1(\theta),...,\mathbf{z}^m(\theta))$ will be injective. Since the various $\mathbf{z}^l$ are the expectations of the normal distribution used, distinct parameter values $\boldsymbol{\theta}$ will then give distinct probability distributions. This is the traditional definition of identifiability of $\boldsymbol{\theta}$ in a MLE setting [23, p. 523], which we see is in accordance with Def. 2 for global identifiability.

3.7. The Fisher information matrix (FIM)

The FIM is a measure of the parameter information in the observations, and we will compute it analytically for a given $\boldsymbol{\theta}$ and design $\{\mathbf{b}^1, ...,\mathbf{b}^m\}$. Let $\mathbf{I}_\theta \in \mathbb{R}^{r_1\times r_1}$ be the FIM for the parameters $\theta \in \mathbb{R}^{r_1}_{>0}$. Entry (j, j') is

$$(\mathbf{I}_\theta)_{j,j'} = \mathbb{E}_Y\left[\left(\frac{\partial \ln f}{\partial\theta_j}\right)\left(\frac{\partial \ln f}{\partial\theta_{j'}}\right)\right], \tag{45}$$

where $\ln f$ is the log-likelihood of Eq. (43). The expectation operator $\mathbb{E}_Y$ is with respect to the density of Eq. (42) for the $\mathbf{Y}^l$s with a fixed $\boldsymbol{\theta}$ [24].

Let $\mathbf{D}^{k\to l}$ be the matrix $\mathbf{D}_{\boldsymbol{\theta}}$ of Eq. (13) where column k is replaced with $\mathbf{c}^l$. Further, let $\mathbf{D}_{(i,j)}$ and $\mathbf{D}^{k\to l}_{(i,j)}$ be the matrices $\mathbf{D}_{\boldsymbol{\theta}}$ and $\mathbf{D}^{k\to l}$ where row i and column j are removed. Finally, let $d_{i,j}$ be entry (i, j) in $\mathbf{D}_{\boldsymbol{\theta}}$. Based on these quantities and Cramer's rule, we are able to calculate $\mathbf{I}_{\boldsymbol{\theta}}$ analytically.

Theorem 5. *Entry (j, j') in $\mathbf{I}_{\boldsymbol{\theta}}$ is given by*

$$(\mathbf{I}_\theta)_{j,j'} = \frac{1}{\sigma^2}\sum_{(l,k)\in\Omega}\left\{\left[\beta_j\left(\frac{\partial z_k^l}{\partial d_{i_j^p,i_j^s}}\right) - \alpha_j\left(\frac{\partial z_k^l}{\partial d_{i_j^s,i_j^s}}\right)\right] \times \left[\beta_{j'}\left(\frac{\partial z_k^l}{\partial d_{i_{j'}^p,i_{j'}^s}}\right) - \alpha_{j'}\left(\frac{\partial z_k^l}{\partial d_{i_{j'}^s,i_{j'}^s}}\right)\right]\right\}, \tag{46}$$

where α_j, β_j, i_j^p and i_j^s are the coefficients and indices of Eq. (2) and z_k^l is the steady-state concentration of X_k under the input vector $\mathbf{b}^l$. The partial derivatives used in Eq. (46) are given by

$$\frac{\partial z_k^l}{\partial d_{i,j}} = (-1)^{i+j}\frac{\det(\mathbf{D}^{k\to l})\det(\mathbf{D}_{(i,j)}) - (1-\delta_{j,k})\det(\mathbf{D}^{k\to l}_{(i,j)})\det(\mathbf{D}_\theta)}{(\det(\mathbf{D}_\theta))^2}. \tag{47}$$

Proof. See Appendix A. □

The $\frac{\partial z_k^l}{\partial d_{i,j}}$'s of Eq. (47) are derivatives of the steady-state concentrations with respect to entries $d_{i,j}$ in the compartmental matrix $\mathbf{D}_{\boldsymbol{\theta}}$. The entries $d_{i,j}$ give the fluxes, and are given by $\boldsymbol{\theta}$ as seen in Eq. (14) for the example.

Note from Eq. (46) that the matrix $\mathbf{I}_{\boldsymbol{\theta}}$ can be written as a sum

$$\mathbf{I}_\theta = \frac{1}{\sigma^2}\sum_{(l,k)\in\Omega}\mathbf{J}_k^l\mathbf{J}_k^{l\,\mathrm{T}} \tag{48}$$

of outer products $\mathbf{J}_k^l\mathbf{J}_k^{l\,\mathrm{T}}$. The vector $\mathbf{J}_k^l \in \mathbb{R}^{r_1}$ is given by

$$\mathbf{J}_k^l = \begin{bmatrix} \beta_1\left(\frac{\partial z_k^l}{\partial d_{i_1^p,i_1^s}}\right) - \alpha_1\left(\frac{\partial z_k^l}{\partial d_{i_1^s,i_1^s}}\right) \\ \vdots \\ \beta_{r_1}\left(\frac{\partial z_k^l}{\partial d_{i_{r_1}^p,i_{r_1}^s}}\right) - \alpha_{r_1}\left(\frac{\partial z_k^l}{\partial d_{i_{r_1}^s,i_{r_1}^s}}\right) \end{bmatrix} = \nabla_\theta z_k^l, \tag{49}$$

and is the gradient of z_k^l with respect to $\boldsymbol{\theta}$ as a column vector. Since each of the outer products in Eq. (48) is a rank-one matrix, $\mathbf{I}_{\boldsymbol{\theta}}$ has rank at most $|\Omega| \le mn$ where $|\Omega|$ is the number of elements in Ω [25]. To obtain a non-singular $\mathbf{I}_{\boldsymbol{\theta}}$, the number $|\Omega|$ of observations must be at least the number r_1 of parameters. Note, however, that an equal or higher number of measurements than parameters is only a necessary condition for identifiability.

The rank of the FIM $\mathbf{I}_{\boldsymbol{\theta}}$ is equal the number of identifiable parameter combinations [18]. It follows that $\boldsymbol{\theta}$ is pointwise identifiable for given values of $\boldsymbol{\theta}$ if and only if $\operatorname{Rank}(\mathbf{I}_\theta) = r_1$, similarly to Thm. 1 for $\mathbf{E}^\theta$. Since $\mathbf{I}_{\boldsymbol{\theta}}$ is a square matrix with dimension r_1, it is then non-singular [26]. We also have that the asymptotic covariance matrix for the estimate $\hat{\theta} = \begin{bmatrix}\hat{\theta}_1 & ... & \hat{\theta}_{r_1}\end{bmatrix}^{\mathrm{T}}$ of $\boldsymbol{\theta}$ in Eq. (44) is given by $\mathbf{I}_\theta^{-1}$ [24]. We will in particular evaluate $\operatorname{Tr}(\mathbf{I}_\theta^{-1})$, giving the asymptotic estimation variance $\sum_{j=1}^{r_1}\operatorname{Var}(\hat{\theta}_j)$.

4. Experimental design

The goal of experimental design is loosely speaking to minimize the parameter estimation uncertainty, see e.g. Fedorov [27] for an introduction. For our experimental setup, there are two things we can control to this end. The first is the choice of the input vectors $\mathbf{b}^l$, and the second is the observation set Ω of Eq. (40) for which metabolites to observe for each input vector. When doing experimental design, a special value for $\boldsymbol{\theta}$ must be chosen. This implies that one should have some a priori knowledge about the value of $\boldsymbol{\theta}$ in order to optimize the experimental design. However, some of the results do not depend on a specific value of $\boldsymbol{\theta}$ and can be applied in the general case.

4.1. Optimal design

We now find the optimal design of input vectors for a fixed value of the parameters $\boldsymbol{\theta}$. An optimal design has to provide pointwise identifiability in accordance with Def. 1, so Thm. 2 should be checked to ensure that this is possible before performing the following optimization procedure. If the standard design gives pointwise identifiability, there exist designs giving pointwise identifiability with $m \ge r_0$ input vectors and potentially also for some $m < r_0$ depending on the network structure. It is also natural to apply Thm. 4 after a design is found to check for global identifiability.

For a fixed number m of input vectors, the D-optimality criterion can be used to find the optimal design. This is done by maximizing the determinant $\det(\mathbf{I}_\theta)$ of the FIM, which is equivalent to maximizing the Shannon information [27]. Note that $\mathbf{I}_{\boldsymbol{\theta}}$ is positive semi-definite [24], as seen from either Eq. (45) or Eq. (48), such that $\det(\mathbf{I}_\theta) > 0$ if $\mathbf{I}_{\boldsymbol{\theta}}$ is non-singular. The D-optimal design of input flux vectors $\mathbf{b}_*^1, ...,\mathbf{b}_*^m \in \mathbb{R}^{r_0}_{\ge 0}$ is the solution of

$$[\mathbf{b}_*^1,...,\mathbf{b}_*^m] = \underset{\mathbf{b}^1,...,\mathbf{b}^m}{\operatorname{argmax}}[\det(\mathbf{I}_\theta)], \tag{50}$$

where $\mathbf{I}_{\boldsymbol{\theta}}$ is given by Eq. (46). The parameter values $\boldsymbol{\theta}$ and the observation set Ω of Eq. (40) are both assumed fixed and given.

This is a maximization in mr_0 variables. Note, however, that if $\mathbf{I}_\theta^{\mathbf{b}}$ is $\mathbf{I}_{\boldsymbol{\theta}}$ for a given design $\{\mathbf{b}^1,...,\mathbf{b}^m\}$, then $\mathbf{I}_\theta^{\gamma\mathbf{b}} = \gamma^2\,\mathbf{I}_\theta^{\mathbf{b}}$ where $\mathbf{I}_\theta^{\gamma\mathbf{b}}$ is $\mathbf{I}_{\boldsymbol{\theta}}$ for the

design $\{\gamma\mathbf{b}^1, ..., \gamma\mathbf{b}^m\}$. This can be seen from Eq. (30) and Eq. (46), and implies that if $\det(\mathbf{I}_\theta^{\mathbf{b}}) > 0$, then $\det(\mathbf{I}_\theta^{\gamma\mathbf{b}}) \to \infty$ for the design $\{\gamma\mathbf{b}^1, ..., \gamma\mathbf{b}^m\}$ as $\gamma \to \infty$. For this reason, Eq. (50) is unbounded for values of m where a design that provides pointwise identifiability exists, and we need constraints on $\{\mathbf{b}^1, ..., \mathbf{b}^m\}$ to get a well-posed problem. A natural constraint is to require the sum of the entries in each vector to be constant such that

$$\|\mathbf{b}^l\|_1 = \sum_{j=1}^{r_0} |b_j^l| = \sum_{j=1}^{r_0} b_j^l = 1 \tag{51}$$

for all $l \in \{1, ..., m\}$. This has the interpretation that the total input flux is equal to one for each of the m input vectors, where the value one is arbitrary since the optimal design only can be determined up to a scaling factor γ as seen above. It is, however, still not guaranteed that there exists a unique solution to the optimization in Eq. (50), which has to be solved numerically.

Note that the ordering of the m vectors $\mathbf{b}_*^l$ is arbitrary, since any permutation of the set $\{\mathbf{b}_*^1, ..., \mathbf{b}_*^m\}$ gives the same value in the maximization of Eq. (50). Also note that the optimal design is independent of the standard deviation since σ can be factored out of $\det(\mathbf{I}_\theta)$ as seen from Eq. (46).

If an optimal addition of one input vector to a given design is wanted, an adjusted formulation of Eq. (50) is given in Appendix D which reduces the maximization problem to a lower dimension.

4.2. Minimal observation set

For a given design $\{\mathbf{b}^1, ..., \mathbf{b}^m\}$, we want to minimize the number $|\Omega|$ of elements in the observation set Ω of Eq. (40) without loosing information. This has similarities to finding minimal sets of output functions needed for identifiability of an ODE system as is done Anguelova et al. [28]. However, we want to find the minimal set with no information loss, not a minimal set that provides identifiability. One could also include the minimization of $|\Omega|$ as a regularization in the optimization of Eq. (50) in the previous subsection.

The FIM $\mathbf{I}_{\boldsymbol{\theta}}$ is additive in the observations as seen in Eq. (46). The term for the observation Y_k^l of the metabolite X_k under $\mathbf{b}^l$ has a factor

$$\frac{\partial z_k^l}{\partial \theta_j} = \beta_j \left(\frac{\partial z_k^l}{\partial d_{i_j^p, i_j^s}} \right) - \alpha_j \left(\frac{\partial z_k^l}{\partial d_{i_j^s, i_j^s}} \right) \tag{52}$$

for all the entries of $\mathbf{I}_{\boldsymbol{\theta}}$ that contain information about the parameter θ_j. Thus, Y_k^l has no information about θ_j if $\frac{\partial z_k^l}{\partial \theta_j} = 0$ such that

$$\beta_j \frac{\partial z_k^l}{\partial d_{i_j^p, i_j^s}} = \alpha_j \frac{\partial z_k^l}{\partial d_{i_j^s, i_j^s}}, \tag{53}$$

since the j'th entry in $\mathbf{J}_k^l$ of Eq. (48) is zero. In particular is Eq. (53) true if

$$\frac{\partial z_k^l}{\partial d_{i_j^p, i_j^s}} = \frac{\partial z_k^l}{\partial d_{i_j^s, i_j^s}} = 0, \tag{54}$$

which corresponds to the steady-state concentration z_k^l of metabolite X_k under input vector $\mathbf{b}^l$ being the same regardless of the value of θ_j. If

$$\frac{\partial z_k^l}{\partial d_{i_j^p, i_j^s}} = \frac{\partial z_k^l}{\partial d_{i_j^s, i_j^s}} = 0 \quad \text{for all } j \in \{1, ..., r_1\}, \tag{55}$$

the measurement Y_k^l does not provide information about any of the parameters since $\mathbf{J}_k^l = \mathbf{0}$ in Eq. (48). This implies that the steady-state concentration z_k^l under $\mathbf{b}^l$ is the same independently of all the parameter values. Under the condition of Eq. (55) we do not loose any information by not observing Y_k^l and can have $(l, k) \notin \Omega$ in accordance with Eq. (40). For this reason, all pairs (l, k) that satisfy Eq. (55) could be excluded from Ω in order to reduce the experimental cost, while the remaining (l, k) should be included in the observation set Ω to not loose any information. The resulting set Ω is minimal in the sense that it is the smallest observation set that provides that same information as the full observation set, but there may exist sets with even fewer elements that still provide identifiability.

The most obvious way to satisfy Eq. (55) for a pair (l, k) is if $z_k^l = 0$, which in Sec. 3.4 was classified as the metabolite X_k being non-active for the input vector $\mathbf{b}^l$. Remember that the classification of active and non-active metabolites does not rely on the value of $\boldsymbol{\theta}$, so the non-active metabolites can be excluded from the observation set Ω for all $\boldsymbol{\theta}$. The active and non-active metabolites are listed in Tab. 4 for the example with the standard design.

5. Example

We will now study the example of Fig. 2 in further detail by finding optimal designs $\{\mathbf{b}_*^1, ..., \mathbf{b}_*^m\}$ and minimal observation sets Ω. Remember that Appendix C provides a simpler example in addition to the one of Fig. 2.

The network of Fig. 2 has $r_1 = 10$ kinetic parameters $\theta = [\theta_1 ... \theta_{10}]^T$ associated with the first order fluxes $\mathbf{v}_1 = [v_1 ... v_{10}]^T$, and $r_0 = 3$ input fluxes $\mathbf{b} = [b_1\ b_2\ b_3]^T$. Recall that $\text{Rank}(\mathbf{S}_P) = r_1 = 10$ for the standard basis design, such that global identifiability for a design is possible by Thm. 3 and 4.

The network has six metabolites X_k, such that for each $\mathbf{b}^l$ we obtain at most six observations Y_k^l. Since the number of parameters is higher than the number of observations for each input vector, it is impossible to obtain identifiability with $m = 1$ input vector. For $m = 2$ vectors we have up to 12 observations such that identifiability of all 10 parameters could be possible.

In the following we will look at pointwise identifiability for the test values

$$\theta_j = 1\ ,\ j = 1, ..., 10 \tag{56}$$

of the parameters. These values are chosen arbitrarily, but much of the discussion is general and independent of the parameter values. The standard deviation factors out of the computations, so a specific σ is not assumed.

The steady-state concentrations $\mathbf{z}^l$ for the standard design $\mathbf{b}^l = \mathbf{e}_l$ and the parameter values of Eq. (56) are given by

$$\begin{aligned}
\mathbf{z}^1 &= -\mathbf{D}_\theta^{-1}\mathbf{S}_0\mathbf{e}_1 = \tfrac{1}{12}[4\ \ 2\ \ 1\ \ 4\ \ 10\ \ 11]^T, \\
\mathbf{z}^2 &= -\mathbf{D}_\theta^{-1}\mathbf{S}_0\mathbf{e}_2 = \tfrac{1}{4}[0\ \ 2\ \ 1\ \ 0\ \ 2\ \ 3]^T \quad \text{and} \\
\mathbf{z}^3 &= -\mathbf{D}_\theta^{-1}\mathbf{S}_0\mathbf{e}_3 = [0\ \ 0\ \ 0\ \ 1\ \ 1\ \ 1]^T
\end{aligned} \tag{57}$$

where $\mathbf{z}$ for an arbitrary input vector $\mathbf{b}$ and the parameters of Eq. (56) is a linear combination of these vectors according to Eq. (30).

5.1. Optimal design

First, we find optimal designs according to the D-optimality criterion of Eq. (50) for the parameter values of Eq. (56) and the constraint of Eq. (51). For now, it is assumed that all metabolites are observed for all the input vectors, i.e. $\Omega = \{(l, k) | 1 \le l \le m, 1 \le k \le n\}$. Based on this, the optimal designs for $m = 2, ..., 6$ are found numerically and shown in Tab. 3 together with the corresponding values of $\det(\mathbf{I}_\theta)$ and $\text{Tr}(\mathbf{I}_\theta^{-1})$. All the designs for $m \ge r_0 = 3$ provide global identifiability by the use of Thm. 4.

As expected we see that the optimal value of $\det(\mathbf{I}_\theta)$ is an increasing function of m since increasing the number of observations increases the information. We also see that with the correct design, $m = 2$ is sufficient for pointwise identifiability. However, this design does not give global identifiability and there is a much larger relative increase in information by increasing m from 2 to 3 than increasing m by 1 when $m \ge 3$. This is probably since there are $r_0 = 3$ degrees of freedom in the choice of each $\mathbf{b}$, not considering the constraint of Eq. (51). Thus, $m \ge 3$ is needed to span the space of possible input flux vectors as seen in

Eq. (29) even though $m = 2$ is sufficient for pointwise identifiability. For this reason, increasing m from 2 to 3 provides completely new information, while increasing m when $m \geq 3$ provides replicates of already existing information. This can also be seen from Thm. 2 where it is shown that the standard design with $m = r_0 = 3$ contains all possible information and Eq. (29) is a part of the proof.

For $m = r_0 = 3$ in Tab. 3 we see that the standard design is optimal. However, this is not a general result as the standard design is in general not optimal. The optimal design for $m \geq 4$ is to keep all the vectors from the $m = 3$ design, and add copies of these for the remaining $\mathbf{b}_*^l$. In particular, note that the optimal for $m = 6$ is two copies of the design for $m = 3$. This shows that the optimal designs for the example have a modular behaviour with modulus $r_0 = 3$, which is reasonable since adding vectors when $m > 3$ repeats already existing information.

5.2. Minimal observation set

We now find the minimal observation set Ω for the standard design, which is optimal for $m = r_0 = 3$ and the parameter values in Eq. (56) as seen in Tab. 3. However, the minimal set will be the same for all parameter values, and we know from Thm. 4 that the standard design gives global identifiability.

Remember that the minimal Ω is all the active metabolites with $z_k^l > 0$ as discussed in Sec. 3.4 and 4.2. This can be found by inspection of Fig. 3 or the steady-state concentrations in Eq. (57), and results in the set of Tab. 4.

We see from Tab. 4 that 5 of the 18 possible observations Y_k^l can be removed without losing information about the 10 parameters θ_j. Remember that the set Ω is minimal in the sense that it is the smallest with the same FIM $\mathbf{I}_{\boldsymbol{\theta}}$ as observing all the concentrations. It may, however, still be possible to find a smaller set that provides global identifiability as the number $|\Omega| = 13$ of elements in Ω is larger than the number $r_1 = 10$ of parameters.

5.3. Global identifiability

We still consider the standard design, which gives global identifiability by Thm. 4. However, we now perform the calculations to show this in details.

First, we look at the information from the $\mathbf{b}^3 = \mathbf{e}_3$ input vector. From the balance equations of Eq. (36) and $\mathbf{S}_1^3$ of Eq. (35) we get the fluxes

$$v_7^3 = v_8^3 = v_{10}^3 = b_3^3 = 1 \tag{58}$$

that are uniquely determined. This implies that θ_7, θ_8 and θ_{10} are found from the input vector $\mathbf{e}_3$, since the flux of a reaction gives the parameter. The three fluxes can then be calculated as $v_j^l = \theta_j z_{i_j}^l$ for the other input vectors.

We continue with the $\mathbf{b}^2 = \mathbf{e}_2$ input vector where we get the fluxes

$$\begin{aligned}
v_4^2 &= b_2^2 - v_8^2 &&= 1 - v_8^2 \\
v_5^2 &= v_8^2 &&= v_8^2 \\
v_6^2 &= -v_8^2 + v_{10}^2 &&= -v_8^2 + v_{10}^2 \\
v_9^2 &= b_2^2 - v_{10}^2 &&= 1 - v_{10}^2
\end{aligned} \tag{59}$$

from the balance equations of Eq. (36) and $\mathbf{S}_1^2$ of Eq. (35). All the fluxes on the right side of Eq. (59) can be calculated from the parameters already found. This gives the values of the four fluxes on the left side, providing the parameters θ_4, θ_5, θ_6 and θ_9 in addition to the three parameters from before.

The three remaining parameters are then θ_1, θ_2 and θ_3, and can be found by applying the $\mathbf{b}^1 = \mathbf{e}_1$ input vector. This gives the three fluxes

$$\begin{aligned}
v_1^1 &= v_4^1 + v_5^1 \\
v_2^1 &= v_7^1 \\
v_3^1 &= -v_5^1 - v_7^1 + v_8^1
\end{aligned} \tag{60}$$

from Eq. (36) and $\mathbf{S}_1^1$ of Eq. (35). All the fluxes on the right side of Eq. (60) can be calculated from the already known parameters, which gives the three fluxes on the left side and the remaining parameters θ_1, θ_2 and θ_3.

The argument above holds for all values of $\boldsymbol{\theta}$, and shows that the standard design provides global identifiability for the example. In reality, information from all the input vectors is used simultaneously. However, the steps above illustrate how the vectors together provide identifiability.

6. Discussion

We have studied identifiability of kinetic parameters under the assumption of first order kinetics and known input fluxes. Theorems for pointwise and global identifiability are presented. The standard basis input vectors are important, and a design can only provide identifiability if the standard one does. Thm. 4 gives a criterion for global identifiability of any given design, and guarantees uniqueness of the estimate $\hat{\boldsymbol{\theta}}$ in Eq. (44) if satisfied. Note that all the identifiability theorems do not depend on a statistical model.

Optimal experimental designs are found by using the Fisher information matrix calculated analytically under the assumption of normal observations. We also create a minimal observation set by removing redundant observations that do not contain any information about the parameters. All the calculations and theorems are illustrated using an example network.

It is assumed that the parameters $\boldsymbol{\theta}$ are constant for all input vectors, but they are likely to vary due to changing external factors such as temperature. To combine data from different experiments and measurements, one could use a mixed model with random effects [29] that allows for variation in $\boldsymbol{\theta}$.

For the calculation of optimal input vectors, a value for $\boldsymbol{\theta}$ has to be assumed. However, this is a standard procedure in experimental design. The minimal observation set, on the other hand, does not rely on the statistical model or the parameter values. It also worth mentioning again that the optimal design is independent of the standard deviation σ of the observations.

The assumption of additive normally distributed error terms with equal variance is easy to adjust to individual variances for the different concentrations. It is also possible to use a multiplicative model like the log-normal distribution. This has only positive support, which is desirable as concentrations by definition are non-negative. Another advantage of the log-normal distribution is that the variance naturally scales with the expectation in contrast to the normal model used here. The potential problem of using a different distribution is the analytical calculation of the Fisher information matrix. However, for the log-normal distribution this can still be done.

The framework of the manuscript can be adopted to other network structures with zero and first order kinetics, including unknown input fluxes and partially known internal fluxes. Also, the global identifiability analysis holds for any kinetic function of one parameter that is bijective in the parameter for a fixed concentration. Altogether, this makes it possible to apply some of the methods in the manuscript to a much larger class of models.

From a practical point of view, it may not be possible to measure all the steady-state concentrations of the network. The observation set Ω of Eq. (40) can then be used to specify which measurements that are available. In some cases one may only have measurements of fluxes or

gene expressions, not of concentrations. The methods of this manuscript can then not be applied directly, but may still give valuable information about the network.

It is also unclear if the input fluxes can be chosen freely, as assumed for the optimal design. However, it could be of interest to evaluate the identifiability of a given design and compare it to the optimal one. Finally, we believe that it is possible that future technological developments will make our methods even more relevant and applicable.

Funding

This work was supported by the Research Council of Norway through grant 248840, dCod 1.0.

Declaration of Competing Interest

We declare that there are no conflicts of interest.

Appendix A. Calculation of the Fisher information matrix

In this section we prove the expression of Eq. (46) for the FIM given in Thm. 5. By the definition in Eq. (45), entry (q, q') of $\mathbf{I}_{\boldsymbol{\theta}}$ is given by

$$(\mathbf{I}_{\theta})_{q,q'} = \mathbb{E}_Y\left[\left(\frac{\partial \ln f}{\partial \theta_q}\right)\left(\frac{\partial \ln f}{\partial \theta_{q'}}\right)\right] \tag{A.1}$$

where the log-likelihood function is given by Eq. (43). Since l and k represents independent observations, the Fisher information matrix will be additive in the observations [24, Ch. 2.6]. This implies that Eq. (A.1) becomes

$$(\mathbf{I}_{\theta})_{q,q'} = \sum_{(l,k)\in\Omega} \mathbb{E}_Y\left[\left(\frac{\partial \ln f_k^l}{\partial \theta_q}\right)\left(\frac{\partial \ln f_k^l}{\partial \theta_{q'}}\right)\right] \tag{A.2}$$

where the marginal log-likelihoods $\ln f_k^l$ are given by the logarithm of Eq. (41).

The derivatives in Eq. (A.2) can then be calculated using the chain rule

$$\frac{\partial \ln f_k^l}{\partial \theta_q} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial \ln f_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial d_{i,j}}{\partial \theta_q}\right) \tag{A.3}$$

with respect to the matrix entries $d_{i,j}$ of $\mathbf{D}_{\boldsymbol{\theta}}$ since the dependence on $\boldsymbol{\theta}$ in $\ln f_k^l$ is through $\mathbf{D}_{\boldsymbol{\theta}}$. For fixed values of $\boldsymbol{\theta}$, all derivatives $\frac{\partial d_{i,j}}{\partial \theta_q}$ in Eq. (A.3) are constants and the entries of Eq. (A.2) then may be written

$$(\mathbf{I}_{\theta})_{q,q'} = \sum_{(l,k)\in\Omega}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i'=1}^{n}\sum_{j'=1}^{n}\left(\frac{\partial d_{i,j}}{\partial \theta_q}\right)\left(\frac{\partial d_{i',j'}}{\partial \theta_{q'}}\right)\mathbb{E}_Y\left[\left(\frac{\partial \ln f_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial \ln f_k^l}{\partial d_{i',j'}}\right)\right] \tag{A.4}$$

which shows that $\mathbf{I}_{\boldsymbol{\theta}}$ is formed by linear combinations of expressions

$$\mathbb{E}_Y\left[\left(\frac{\partial \ln f_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial \ln f_k^l}{\partial d_{i',j'}}\right)\right]. \tag{A.5}$$

The derivatives in the expression of Eq. (A.5) are given by

$$\frac{\partial \ln f_k^l}{\partial d_{i,j}} = \frac{1}{\sigma^2}(Y_k^l - z_k^l)\left(\frac{\partial z_k^l}{\partial d_{i,j}}\right) \tag{A.6}$$

where the derivative is fixed for a given $\boldsymbol{\theta}$ such that Eq. (A.5) becomes

$$\begin{aligned}\mathbb{E}_Y\left[\left(\frac{\partial \ln f}{\partial d_{i,j}}\right)\left(\frac{\partial \ln f}{\partial d_{i',j'}}\right)\right] &= \frac{1}{\sigma^4}\left(\frac{\partial z_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial z_k^l}{\partial d_{i',j'}}\right)\mathbb{E}_Y[(Y_k^l - z_k^l)^2] \\ &= \frac{1}{\sigma^4}\left(\frac{\partial z_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial z_k^l}{\partial d_{i',j'}}\right)\mathrm{Var}(Y_k^l) = \frac{1}{\sigma^2}\left(\frac{\partial z_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial z_k^l}{\partial d_{i',j'}}\right).\end{aligned} \tag{A.7}$$

Plugging this into Eq. (A.4) gives

$$(\mathbf{I}_{\theta})_{q,q'} = \frac{1}{\sigma^2}\sum_{(l,k)\in\Omega}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i'=1}^{n}\sum_{j'=1}^{n}\left[\left(\frac{\partial d_{i,j}}{\partial \theta_q}\right)\left(\frac{\partial d_{i',j'}}{\partial \theta_{q'}}\right)\left(\frac{\partial z_k^l}{\partial d_{i,j}}\right)\left(\frac{\partial z_k^l}{\partial d_{i',j'}}\right)\right] \tag{A.8}$$

for the entries of $\mathbf{I}_{\boldsymbol{\theta}}$. The expression of Eq. (A.8) may seem overwhelming with several sums. However, the only assumption used is that the entries of $\mathbf{D}_{\boldsymbol{\theta}}$ are functions of the parameters $\boldsymbol{\theta}$, and four of the sums will collapse by making use of the structure of $\mathbf{D}_{\boldsymbol{\theta}}$. First, we refactor Eq. (A.8) as

$$(\mathbf{I}_{\theta})_{q,q'} = \frac{1}{\sigma^2}\sum_{(l,k)\in\Omega}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial d_{i,j}}{\partial \theta_q}\right)\left(\frac{\partial z_k^l}{\partial d_{i,j}}\right)\right]\left[\sum_{i'=1}^{n}\sum_{j'=1}^{n}\left(\frac{\partial d_{i',j'}}{\partial \theta_{q'}}\right)\left(\frac{\partial z_k^l}{\partial d_{i',j'}}\right)\right] \tag{A.9}$$

such that the two brackets are associated with θ_q and $\theta_{q'}$, respectively. The matrix $\mathbf{D}_{\theta} = \{d_{i,j}\}$ of Eq. (13) is given by $\mathbf{D}_{\theta} = \mathbf{S}_1\mathbf{K}_{\theta}$. The only occurrence of θ_q in $\mathbf{K}_{\boldsymbol{\theta}}$ is in position (q, i_q^s). Then, only column i_q^s in $\mathbf{D}_{\boldsymbol{\theta}}$ contains θ_q. Thus, the sums over j and j' in Eq. (A.9) reduce to only $j = i_q^s$ and $j' = i_{q'}^s$. Further, since $\mathbf{K}_{\boldsymbol{\theta}}$ is exactly equal to θ_q in position (q, i_q^s) we have the derivatives

$$\frac{\partial d_{i,i_q^s}}{\partial \theta_q} = (\mathbf{S}_1)_{i,q} \tag{A.10}$$

where $(\mathbf{S}_1)_{i,q}$ is entry $(i,\ q)$ of $\mathbf{S}_1$. Then Eq. (A.9) simplifies to

$$(\mathbf{I}_\theta)_{q,q'} = \frac{1}{\sigma^2} \sum_{(l,k)\in\Omega} \left[\sum_{i=1}^{n} (\mathbf{S}_1)_{i,q} \left(\frac{\partial z_k^l}{\partial d_{i,i_q^s}} \right) \right] \left[\sum_{i'=1}^{n} (\mathbf{S}_1)_{i',q'} \left(\frac{\partial z_k^l}{\partial d_{i',i_{q'}^s}} \right) \right]. \tag{A.11}$$

Each column q in $\mathbf{S}_1$ has two non-zero entries, namely $-\alpha_q$ in row i_q^s and β_q in row i_q^p. The sums over i and i' in Eq. (A.11) then reduce such that

$$\begin{aligned}(\mathbf{I}_\theta)_{q,q'} = \frac{1}{\sigma^2} \sum_{(l,k)\in\Omega} &\left\{ \left[\beta_q \left(\frac{\partial z_k^l}{\partial d_{i_q^p,i_q^s}} \right) - \alpha_q \left(\frac{\partial z_k^l}{\partial d_{i_q^s,i_q^s}} \right) \right] \right. \\ &\left. \times \left[\beta_{q'} \left(\frac{\partial z_k^l}{\partial d_{i_{q'}^p,i_{q'}^s}} \right) - \alpha_{q'} \left(\frac{\partial z_k^l}{\partial d_{i_{q'}^s,i_{q'}^s}} \right) \right] \right\}.\end{aligned} \tag{A.12}$$

By changing the notation to j and j' instead of q and q', we have Eq. (46) of Thm. 5 and see that we are left with at most four non-zero terms for each pair of l and k independently of the network size.

The remaining part is to calculate the partial derivatives used in Eq. (46) where the various matrices used are explained in relation to Thm. 5. By Eq. (17) and Cramer's rule, z_k^l can be expressed as

$$z_k^l = -\frac{\det(\mathbf{D}^{k\to l})}{\det(\mathbf{D}_\theta)}. \tag{A.13}$$

To prepare for differentiation with respect to $d_{i,j}$, one may write

$$\det(\mathbf{D}_\theta) = (-1)^{i+j} d_{i,j} \det(\mathbf{D}_{(i,j)}) + (\text{without } d_{i,j}) \quad \text{and} \tag{A.14}$$

$$\det(\mathbf{D}^{k\to l}) = \begin{cases} (\text{without } d_{i,j}) & \text{if } j = k \\ (-1)^{i+j} d_{i,j} \det(\mathbf{D}_{(i,j)}^{k\to l}) + (\text{without } d_{i,j}) & \text{if } j \neq k \end{cases} \tag{A.15}$$

which gives the partial derivatives

$$\frac{\partial}{\partial d_{i,j}} \det(\mathbf{D}_\theta) = (-1)^{i+j} \det(\mathbf{D}_{(i,j)}) \quad \text{and} \tag{A.16}$$

$$\begin{aligned}\frac{\partial}{\partial d_{i,j}} \det(\mathbf{D}^{k\to l}) &= \begin{cases} 0 & \text{if } j = k \\ (-1)^{i+j} \det(\mathbf{D}_{(i,j)}^{k\to l}) & \text{if } j \neq k \end{cases} \\ &= (1-\delta_{j,k})(-1)^{i+j} \det(\mathbf{D}_{(i,j)}^{k\to l}).\end{aligned} \tag{A.17}$$

By applying the quotient rule to Eq. (A.13) we then get

$$\frac{\partial z_k^l}{\partial d_{i,j}} = (-1)^{i+j} \frac{\det(\mathbf{D}^{k\to l})\det(\mathbf{D}_{(i,j)}) - (1-\delta_{j,k})\det(\mathbf{D}_{(i,j)}^{k\to l})\det(\mathbf{D}_\theta)}{(\det(\mathbf{D}_\theta))^2} \tag{47}$$

which is Eq. (47) of Thm. 5.

Appendix B. Identifiability examples

To illustrate the identifiability further, we have a few more examples for the network of Fig. 2 similar to the one given after Def. 1. First, we look closer at the design $\{\mathbf{e}_2, \mathbf{e}_3\}$. From the argument in Sec. 5.3, all the parameters except θ_1, θ_2 and θ_3 are identifiable for this design independently of the values of $\boldsymbol{\theta}$. To show this, we solve Eq. (28) for $\boldsymbol{\theta}$ with the parameter values of Eq. (56) and look at the pointwise identifiability. The solution is

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \\ \theta_7 \\ \theta_8 \\ \theta_9 \\ \theta_{10} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \gamma_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \gamma_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \gamma_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{B.1}$$

where $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ are free parameters, and the constant vector is the true parameter values of Eq. (56). We see that the three parameters θ_1, θ_2 and θ_3 are free as proposed. They do not take part in linear combinations either, as we have absolutely no information about them from the observations.

If we instead look at the design $\{\mathbf{e}_1, \mathbf{e}_3\}$ for the parameter values in Eq. (56), we get a different situation. The solution of Eq. (28) is then

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \\ \theta_7 \\ \theta_8 \\ \theta_9 \\ \theta_{10} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{B.2}$$

where $\gamma \in \mathbb{R}$ again is a free parameter and the constant vector is the true parameter values. There are still three non-identifiable parameters, but only one free parameter in the solution. Looking at Fig. 2, we see that the non-identifiable parameters θ_1, θ_3 and θ_5 form a circuit. By using the input vectors $\mathbf{e}_1$ and $\mathbf{e}_3$ we are not able to determine how much of the flow that goes through v_1 and v_5, and how much of the flow that goes through v_3.

Appendix C. Toy example

In this appendix we illustrate the method of the manuscript by applying it to the small toy example in Fig. C.4, which is a sub-network of the example in Fig. 1. It has $n = 2$ metabolites X_k, $r_1 = 3$ parameters θ_j and $r_0 = 2$ input fluxes b_j where all the $r = 5$ reactions are given in Tab. C.5.

We see that the reactions are already numbered according to Eq. (5) with the first order fluxes first and the leaving fluxes last. The stoichiometric matrix $\mathbf{S}$ is then given by the two sub-matrices $\mathbf{S}_1$ and $\mathbf{S}_0$ as

$$\mathbf{S} = [\mathbf{S}_1 \;\; \mathbf{S}_0] = \begin{bmatrix} -1 & -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 1 \end{bmatrix} \quad \text{where} \tag{C.1}$$

$$\mathbf{S}_1 = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{C.2}$$

The corresponding separation of the fluxes $\mathbf{v}$ into $\mathbf{v}_1$ and $\mathbf{v}_0$ is given by

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_0 \end{bmatrix} \quad \text{where} \quad \mathbf{v}_1 = \begin{bmatrix} \theta_1 x_1 \\ \theta_2 x_1 \\ \theta_3 x_2 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_0 = \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{C.3}$$

We now want to check for global identifiability in accordance with Def. 2 using Thm. 3 and 4. To do this, we consider the design $\{\mathbf{b}^1, \mathbf{b}^2\} = \{\mathbf{e}_1, \mathbf{e}_2\}$ of standard basis vectors $\mathbf{e}_1 = [1 \;\; 0]^T$ and $\mathbf{e}_2 = [0 \;\; 1]^T$ in $\mathbb{R}^2$. The active reactions and metabolites for this design are given in Tab. C.6 and visualized in Fig. C.5, where the sets are easily found by a visual inspection of Fig. C.4.

From $\mathbf{S}_1$ in Eq. (C.2) and the active reactions in Tab. C.6 we get

$$\mathbf{S}_1^1 = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_1^2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & -1 \end{bmatrix} \tag{C.4}$$

as the modified stoichiometric matrices $\mathbf{S}_1^l$ of Sec. 3.4 for the standard design where the gray columns of non-active reactions are set to zero. This gives

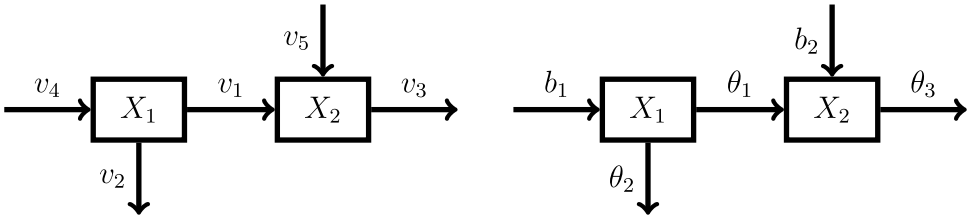

Fig. C1. Toy example network with two different sets of labels. The nodes are metabolites with names X_k and concentrations x_k, while the directed edges are irreversible reactions. In the left panel are all the fluxes labelled by v_j. In the right panel are the first order fluxes labelled by their parameters θ_j, while the input fluxes are labelled by b_j.

Table C1
Reactions for the example in Fig. C.4 on the form of Eq. (2).

$j = 1$	X_1	$\rightarrow$	X_2	$i_1^s = 1$	$i_1^p = 2$
$j = 2$	X_1	$\rightarrow$		$i_2^s = 1$	
$j = 3$	X_2	$\rightarrow$		$i_3^s = 2$	
$j = 4$		$\rightarrow$	X_1		$i_4^p = 1$
$j = 5$		$\rightarrow$	X_2		$i_5^p = 2$

Table C2
Active reactions (left) and metabolites (right) for the example in Fig. C.4 and the standard vectors $\mathbf{e}_j$, where a cross is active. The information is visualized in Fig. C.5.

	v_1	v_2	v_3		X_1	X_2
$\mathbf{e}_1$	×	×	×	$\mathbf{e}_1$	×	×
$\mathbf{e}_2$			×	$\mathbf{e}_2$		×

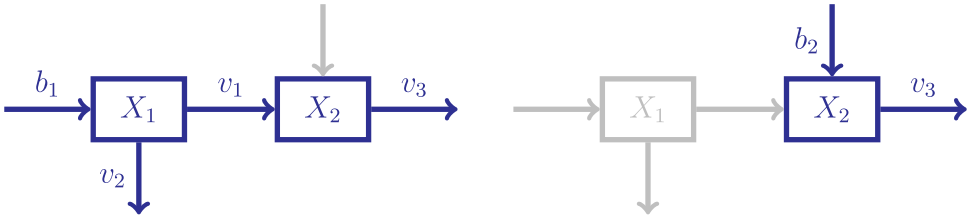

Fig. C2. Active reactions and metabolites for the example in Fig. C.4 and the standard design, same as in Tab. C.6. Dark blue fluxes and metabolites are active, while the light gray ones are non-active. The left panel is for $\mathbf{b}^1 = \mathbf{e}_1$ and the right panel for $\mathbf{b}^2 = \mathbf{e}_2$.

$$\mathbf{S}_P = \begin{bmatrix} \mathbf{S}_1^1 \\ \mathbf{S}_1^2 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{C.5}$$

for the matrix $\mathbf{S}_P$ of Eq. (37). Since $\mathbf{S}_P$ has $\text{Rank}(\mathbf{S}_P) = r_1 = 3$ giving full column rank, the parameter vector $\boldsymbol{\theta}$ is globally identifiable by Thm. 4.

The global identifiability can be also be shown by similar calculations as in Sec. 5.3 for the example of Fig. 1. As before, let z_k^l and v_j^l be the steady-state values of x_k and v_j for the input vector $\mathbf{b}^l = \mathbf{e}_l$. From the steady-state condition of Eq. (36) we then get the balance equations

$$\text{For } X_1 \text{ under } \mathbf{e}_1\text{:} 1 = b_1^1 = v_1^1 + v_2^1 \tag{C.6}$$

$$\text{For } X_2 \text{ under } \mathbf{e}_1\text{:} v_1^1 = v_3^1 \tag{C.7}$$

$$\text{For } X_2 \text{ under } \mathbf{e}_2\text{:} 1 = b_2^2 = v_3^2 \tag{C.8}$$

where the fluxes are given by $v_j^l = \theta_j z_{i_j^s}^l$ as before with the substrate indices i_j^s given in Tab. C.5 and all the steady-state concentrations z_k^l assumed known.

By using that $v_3^2 = \theta_3 z_2^2$, Eq. (C.8) gives us the parameter value $\theta_3 = \frac{1}{z_2^2}$. Then $v_3^1 = \theta_3 z_2^1 = \frac{z_2^1}{z_2^2}$ such that Eq. (C.7) gives $\theta_1 = \frac{z_2^1}{z_1^1 z_2^2}$ since $v_1^1 = \theta_1 z_1^1$. Finally, from using that $v_2^1 = \theta_2 z_1^1$ in Eq. (C.6) we get $\theta_2 = \frac{z_2^2 - z_2^1}{z_1^1 z_2^2}$. Note that $z_2^2 > z_2^1$ such that $\theta_2 > 0$ since $1 = v_3^2 = \theta_3 z_2^2 > v_3^1 = \theta_3 z_2^1 = 1 - v_2^1$. In total, we see that all the parameters $\boldsymbol{\theta}$ can be calculated from the three concentrations z_1^1, z_2^1 and z_2^2 for the standard design such that $\boldsymbol{\theta}$ is globally identifiable. The three concentrations are exactly the active metabolites of Tab. C.6 which gives the minimal observation set

$$\Omega = \{(1, 1), (1, 2), (2, 2)\} \tag{C.9}$$

of Sec. 4.2 where $(l, k) \in \Omega$ if $z_k^l > 0$ such that X_k is active for $\mathbf{b}^l = \mathbf{e}_l$. It could also be noted that if all $\theta_j = 1$, the optimal design of Eq. (50) with $m = 2$ vectors under the constraint of Eq. (51) is simply the standard design. The corresponding maximum value for Eq. (50) is $\det(\mathbf{I}_\theta) = 0.0156 \cdot \sigma^{-6}$.

Appendix D. Addition of an input vector to an existing design

The maximization of $\det(\mathbf{I}_\theta)$ in Eq. (50) under the constraint of Eq. (51) gives a criterion for a design of m input vectors to be optimal. In general, the full maximization in Eq. (50) must be repeated if the optimal design of $m + 1$ input vectors is wanted instead. However, in a practical situation one may have a fixed design of m input vectors and want to add one more input vector in an optimal way. If the fixed design of m input vectors provides pointwise identifiability, the maximization of Eq. (50) can be adopted to this situation. Note that the fixed design is not assumed to be optimal, only to provide pointwise identifiability.

Theorem 6. *Assume that a design* $\{\mathbf{b}^1, ..., \mathbf{b}^m\}$ *provides pointwise identifiability, and that we want to add one more input vector to the design in an optimal way. When keeping the m original input vectors fixed and assuming that* $(m + 1, k) \in \Omega$ *for all k, the optimization*

$$\mathbf{b}_*^{m+1} = \underset{\mathbf{b}^{m+1}}{\text{argmax}}[\det(\mathbf{I}_\theta)] \tag{D.1}$$

for the additional input vector is equivalent to

$$\mathbf{b}_*^{m+1} = \underset{\mathbf{b}^{m+1}}{\text{argmax}}\left[\det\left(\mathbb{I}_n + \frac{1}{\sigma^2}\mathbf{J}^{m+1\text{T}}(\mathbf{I}_\theta^m)^{-1}\mathbf{J}^{m+1}\right)\right], \tag{D.2}$$

where $\mathbf{I}_\theta^m$ *is the FIM for the original design,* $\mathbb{I}_n$ *is the* $n \times n$ *identity matrix and* $\mathbf{J}^{m+1} = [\mathbf{J}_1^{m+1} \ \ ... \ \ \mathbf{J}_n^{m+1}]$ *with the vectors* $\mathbf{J}_k^{m+1}$ *given by Eq. (49).*

Proof. The FIM $\mathbf{I}_\theta$ for the new design of $m + 1$ input vectors can be written

$$\mathbf{I}_\theta = \frac{1}{\sigma^2}\sum_{\substack{l=1 \\ (l,k)\in\Omega}}^{m+1}\sum_{k=1}^{n}\mathbf{J}_k^l\mathbf{J}_k^{l\text{T}} = \mathbf{I}_\theta^m + \frac{1}{\sigma^2}\mathbf{J}^{m+1}\mathbf{J}^{m+1\text{T}} \tag{D.3}$$

by using Eq. (48) where the vectors $\mathbf{J}_k^l$ are given by Eq. (49), and Ω is the observation set for the full design of $m + 1$ input vectors. Since the original

design of m input vectors provides identifiability, the matrix $\mathbf{I}_\theta^m$ is invertible and symmetric positive definite with $\det(\mathbf{I}_\theta^m) > 0$. Then

$$\begin{aligned}\det(\mathbf{I}_\theta) &= \det\left(\mathbf{I}_\theta^m + \frac{1}{\sigma^2}\mathbf{J}^{m+1}\mathbf{J}^{m+1\mathrm{T}}\right)\\ &= \det\left(\mathbb{I}_n + \frac{1}{\sigma^2}\mathbf{J}^{m+1\mathrm{T}}(\mathbf{I}_\theta^m)^{-1}\mathbf{J}^{m+1}\right)\det(\mathbf{I}_\theta^m)\end{aligned} \tag{D.4}$$

by the matrix determinant lemma [30]. Since $\det(\mathbf{I}_\theta^m)$ is constant with respect to the optimization of Eq. (D.1) and has a positive value, the optimum for $\det(\mathbf{I}_\theta)$ can be found by considering only the first factor of Eq. (D.4). □

This theorem transforms the optimization from evaluating a determinant of dimension r_1, the number of parameters and first order reactions, to evaluating a determinant of dimension n, the number of metabolites. Since $r_1 > n$ in most networks, this may be beneficial even though the formula of Eq. (D.2) involves a fixed matrix inverse $(\mathbf{I}_\theta^m)^{-1}$.

Note that the optimization of Eq. (D.2) is unbounded in the same way as Eq. (50), so a constraint like Eq. (51) is still necessary for the additional input vector. It is also worth mentioning that the method for the extra input vector easily can be adjusted if $(m+1, k) \notin \Omega$ for some k.

References

[1] W.L. Oberkampf, C.J. Roy, Verification and Validation in Scientific Computing, Cambridge University Press, 2010.
[2] C.J. Roy, W.L. Oberkampf, A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing, Comput. Methods Appl. Mech. Eng. 200 (2011) 2131–2144.
[3] H. Miao, X. Xia, A.S. Perelson, H. Wu, On identifiability of nonlinear ode models and applications in viral dynamics, SIAM Rev. 53 (2011) 3–39.
[4] A. Raue, J. Karlsson, M.P. Saccomani, M. Jirstrand, J. Timmer, Comparison of approaches for parameter identifiability analysis of biological systems, Bioinformatics 30 (10) (2014) 1440–1448.
[5] S.M. Baker, C.H. Poskar, F. Schreiber, B.H. Junker, A unified framework for estimating parameters of kinetic biological models, BMC Bioinf. 16 (104) (2015).
[6] O.-T. Chis, J.R. Banga, E. Balsa-Canto, Structural identifiability of systems biology models: a critical comparison of methods, PloS one 6 (11) (2011).
[7] K.A. McLean, K.B. McAuley, Mathematical modelling of chemical processesâobtaining the best model predictions and parameter estimates using identifiability and estimability procedures, Can. J. Chem. Eng. 90 (2012) 351–366.
[8] C. Kravaris, J. Hahn, Y. Chu, Advances and selected recent developments in state and parameter estimation, Comput. Chem. Eng. 51 (2013) 111–123.
[9] J. Vanlier, C. Tiemann, P. Hilbers, N. Van Riel, Parameter uncertainty in biochemical models described by ordinary differential equations, Math. Biosci. 246 (2013) 305–314.
[10] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics 25 (2009) 1923–1929.
[11] C. Kreutz, An easy and efficient approach for testing identifiability, Bioinformatics 34 (2018) 1913–1921.
[12] Y. Zuo, S. Sarkar, M.T. Corwin, K.A. Olson, R.D. Badawi, G. Wang, Structural and practical identifiability of dual-input kinetic modeling in dynamic pet of liver inflammation, Phys. Med. Biol. 64 (17) (2019).
[13] A. Gábor, A.F. Villaverde, J.R. Banga, Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems, BMC Syst. Biol. 11 (54) (2017).
[14] D. Casas-Orozco, A.L. Villa, O.J. Guerra, G.V. Reklaitis, Dynamic parameter estimation and identifiability analysis for heterogeneously-catalyzed reactions: catalytic synthesis of nopol, Chem. Eng. Res. Des. 134 (2018) 226–237.
[15] S. Berthoumieux, M. Brilli, D. Kahn, H. De Jong, E. Cinquemani, On the identifiability of metabolic network models, J. Mathem. Biol. 67 (2013) 1795–1832.
[16] D. Anderson, Compartmental Modeling and Tracer Kinetics, Lecture Notes in Biomathematics, Springer, 1983.
[17] W. Wiechert, 13C metabolic flux analysis, Metab. Eng. 3 (2001) 195–206.
[18] M.C. Eisenberg, M.A. Hayashi, Determining identifiable parameter combinations using subset profiling, Math. Biosci. 256 (2014) 116–126.
[19] C. Conradi, E. Feliu, M. Mincheva, C. Wiuf, Identifying parameter regions for multistationarity, PLoS Comput. Biol. 13 (10) (2017).
[20] B.Ø. Palsson, Systems Biology: Constraint-based Reconstruction and Analysis, Cambridge University Press, 2015.
[21] B. Ingalls, Mathematical Modelling in Systems Biology: An introduction, University of Waterloo, 2015. Version December 23, 2015
[22] D.M. Bates, D.G. Watts, Nonlinear Regression Analysis and Its Applications, Wiley and Sons: New York, 1988.
[23] G. Casella, R. Berger, Statistical Inference, Duxbury advanced series in statistics and decision sciences, Thomson Learning, 2002.
[24] E. Lehmann, G. Casella, Theory of Point Estimation, Springer Texts in Statistics, Springer New York, 2006.
[25] L. Trefethen, D. Bau, Numerical Linear Algebra, Other Titles in Applied Mathematics, Society for Industrial and Applied Mathematics, 1997.
[26] S.D. Silvey, Statistical Inference, Monographs on applied probability and statistics, Chapman and Hall, 1975.
[27] V. Fedorov, Theory Of Optimal Experiments, Probability and Mathematical Statistics, Elsevier Science, 1972.
[28] M. Anguelova, J. Karlsson, M. Jirstrand, Minimal output sets for identifiability, Math. Biosci. 239 (2012) 139–153.
[29] D.L. Janzén, M. Jirstrand, M.J. Chappell, N.D. Evans, Extending existing structural identifiability analysis methods to mixed-effects models, Math. Biosci. 295 (2018) 1–10.
[30] D.A. Harville, Matrix Algebra from a Statistician's Perspective, 1 Springer, 1997.