Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-2 (2023) 581–586

Indistinguishability analysis and observer design for size-structured cell populations

P. Jerono * **J. J. Winkin** ** **A. Vande Wouwer** *** **A. Schaum** ****

* *Digital Process Engineering Group (DPE), Karlsruhe Institute of Technology, Germany, pascal.jerono@kit.edu*

** *Department of Mathematics and Institute for Complex Systems (naXys), University of Namur, Belgium*

*** *Systems, Estimation, Control and Optimization (SECO), University of Mons, Belgium*

**** *Chair of Automation and Control (ACON), Kiel University, Germany*

Abstract: The observability property of population balance equations with biomass measurements is addressed within the framework of indistinguishable trajectories. The population balance equation is described by a partial integro-differential equation which is coupled with an ordinary differential equation for the nutrient dynamics. For the semi-discretized model equations, the dynamics of indistinguishable trajectories are derived and evaluated for the special case of equal partitioning at cell division revealing that the observability property is guaranteed as long as the nutrient concentration is not depleted. Based on these results, an observer is designed and tested in simulation to estimate the cell population distribution using biomass measurements in a batch bioreactor for the case of a bi-structured population.

Keywords: Observability, indistinguishable trajectories, cell population dynamics

1. INTRODUCTION

In bioprocess technology, the lack of available online measurement information is a common issue due to physical constraints of the reactor or limited sensor availability. The design of suitable observers to estimate relevant process variables based on available measurement data is therefore of great interest and is a key enabler of robust control strategies. The convergence of the estimated observer states to the real system states is associated with the system's observability property and its outputs (measurements). The analysis of this property usually relies on calculating time derivatives of the outputs, which is related to different levels of complexity depending on the considered system and output properties. For nonlinear systems, the local observability can be investigated by exploiting the linearization of the observability map (Zeitz, 1990; Isidori et al., 1995). For higher dimensional systems, the analysis can be carried out by a graph analytical approach investigating the structural observability, i.e. whether the observability property is satisfied without taking any parameter values into account (Liu et al., 2013; Jerono et al., 2021a). Nevertheless, local and structural observability are weaker properties that can not guarantee convergence of the observer in every scenario. A more rigorous investigation of the (global) observability property is given by the analysis of the indistinguishable trajectories (Ibarra et al., 2004; Moreno and Dochain, 2005; Schaum and Moreno, 2007; Moreno et al., 2014). In this method, the dynamics of indistinguishable trajectories, i.e. those leading to the same system input-output behavior, are analyzed. The observability property is then concluded by the analysis of these trajectories, leading to more model insight. As a trade–off for investigating the observability property via indistinguishable trajectories, the analysis is usually more involved as it considers a dynamical system rather than an algebraic property.

In this work, the observability analysis of a cell population balance model is carried out within the framework of indistinguishable trajectories. Cell population balance models are usually described by a partial integro-differential equation which is coupled to a set of nonlinear ordinary differential equations (ODEs) to describe the nutrient dynamics (Tsuchiya et al., 1966; Mantzaris et al., 1999). Following an early lumping approach, the resulting system is given by a set of nonlinear differential equations with the system dimension depending on the number of discretization points. The structural observability property for the considered system class is investigated in (Jerono et al., 2021a). In the present work, the indistinguishable trajectories of the system with respect to biomass measurements are analyzed. The dynamics of indistinguishable trajectories are derived for the general model equations and investigated in more detail for the special case of equal partitioning of the cells at division. Further, an observer is designed to estimate a bi–partitioned cell distribution based on biomass measurements in a batch reactor and tested in simulation.

⋆ This work is supported by funding from the naXys institute of the University of Namur.

10.1016/j.ifacol.2023.10.1630

2. PROBLEM FORMULATION

Consider the cell population balance equation

$$\partial_t n(m,t) = -\partial_m[r(m,s)n(m,t)] - \Gamma(m,s)n(m,t) + 2\int_m^{m^*} \Gamma(\mu,s)p(m,\mu)n(\mu,t)\mathrm{d}\mu \tag{1a}$$

$$n(m^*,t) = 0, \quad n(m,0) = n_0(m) \tag{1b}$$

where $t \in \mathbb{R}_{\geq 0}$ is the time, $m \in [m_*, m^*]$ is the cell mass, m_* the minimum and m^* the maximum mass of a cell and $n(m,t) : [m_*, m^*] \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is the cell mass distribution density function. The cell growth rate is denoted as $r(m,s) : [m_*, m^*] \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with the substrate concentration $s \in \mathbb{R}_{\geq 0}$. The cell division rate is given by $\Gamma(m,s) : [m_*, m^*] \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $p(m,\mu) : [m_*, m^*] \times [m_*, m^*] \to \mathbb{R}_{\geq 0}$ is the division probability density function which determines the possibility that a mother cell of mass μ divides into daughter cells of mass m. Note that for the considered model (1) the existence and positivity of the solutions in $L^1 \times \mathbb{R}_{\geq 0}$ is shown in (Beniich et al., 2018). To take into account the substrate dynamics, the cell population balance equation is coupled with

$$\dot{s} = -\int_{m_*}^{m^*} r(m,s)n(m,t)\mathrm{d}m, \tag{1c}$$

so that the model given by (1) and (1c) consists of a partial integro-differential equation coupled with a differential equation accounting for the substrate dynamics. Further, it is considered that growth is proportional to mass and division is proportional to growth, i.e.

$$r(m,s) = \rho(s)m, \qquad \Gamma(m,s) = \gamma(m)\rho(s)$$

and $\rho(s)$ is given by the Monod growth rate

$$\rho(s) = k_s \frac{s}{K_s + s}. \tag{2}$$

Recalling mass conservation (Mantzaris and Daoutidis, 2004; Schaum and Jerono, 2019), i.e.

$$\int_{m_*}^{m^*} m\Gamma(m,s)n(m,t)\mathrm{d}m = \int_{m_*}^{m^*} 2m\left(\int_{m_*}^{m^*} \Gamma(\mu,s)p(m,\mu)n(m,t)\mathrm{d}\mu\right)\mathrm{d}m \tag{3}$$

and calculating the first moment of the cell population balance equation (1a)

$$b = \int_{m_*}^{m^*} mn(m,t)\mathrm{d}m \tag{4}$$

where b denotes the biomass, one has

$$\dot{b} = \rho(s)\int_{m_*}^{m^*} mn(m,t)\mathrm{d}m = \rho(s)b. \tag{5}$$

Note that in terms of the total biomass dynamics the model equations (1) can be written as

$$\dot{b} = \rho(s)b \tag{6a}$$

$$\dot{s} = -\rho(s)b, \tag{6b}$$

so that the solution of the biomass in the mass balance model (6) is equivalent to the solution of the first moment of the cell distribution (1).

3. DISCRETIZATION

The model equations (1) are discretized in the mass domain, leading to a set of nonlinear ordinary differential equations. A first-order upwind finite difference scheme is used to approximate the partial derivative and the trapezoidal rule is used to approximate the integral term of cell birth and the first moment of the distribution. The discretized model equations then read

$$\dot{n}_i = -\frac{1}{\Delta_m}\rho(s)(m_i n_i - m_{i-1}n_{i-1}) - \Gamma(m_i,s)n_i + 2\Delta_m \sum_{j=i+1}^{z} \Gamma(m_j,s)p(m_i,m_j)n_j \tag{7a}$$

$$\dot{s} = -\rho(s)b = -\rho(s)\Delta_m \sum_{i=1}^{z} m_i n_i =: f_s(\boldsymbol{x}) \tag{7b}$$

$$n_{z+1} = 0, \quad n_i(0) = n_{i,0}, \quad s(0) = s_0 \tag{7c}$$

where n_i is the cell density of mass m_i at time t, the discretization step size is given by Δ_m and z is the number of interior discretization points. Introducing the state vector $\boldsymbol{x} = [\boldsymbol{n}^T, s]^T = [n_1, \ldots, n_z, s]^T$ equations (7) can be re–cast into the form

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{A}(s)\boldsymbol{n} \\ f_s(\boldsymbol{x}) \end{bmatrix}, \ \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^{z+1}_{\geq 0} \tag{8}$$

where $\boldsymbol{A}(s)$ is constructed from (7a) and has a triangular matrix structure with additional elements on the first lower off diagonal

$$\boldsymbol{A}(s) = \rho(s)\boldsymbol{\mathcal{A}} = \rho(s)\begin{bmatrix} * & * & * & * & \cdots & * & * & * \\ * & * & * & * & \cdots & * & * & * \\ 0 & * & * & * & \cdots & * & * & * \\ 0 & 0 & * & * & \cdots & * & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & * & * \end{bmatrix}. \tag{9}$$

Note that the boundaries of the cell distribution n_0 and n_{z+1} can be excluded from the state vector, because of the containment conditions $\rho(s)m_* = 0$ and $n_{z+1} = 0$. In the following, the solution of (8) will be denoted as $\boldsymbol{x}(t) = \boldsymbol{\phi}(t;\boldsymbol{x}_0)$ with $\boldsymbol{\phi} : [0,\infty) \to \mathbb{R}^{z+1}_{\geq 0}$.

4. INDISTINGUISHABLE TRAJECTORIES

The analysis of indistinguishable trajectories addresses the question of whether two trajectories of a system, starting at different initial conditions, can have the same output for all times (Nijmeijer and Van der Schaft, 1990; Ibarra et al., 2004).

Definition 4.1. Two trajectories $\boldsymbol{x}_1(t) = \boldsymbol{\phi}(t;\boldsymbol{x}_{1,0})$ and $\boldsymbol{x}_2(t) = \boldsymbol{\phi}(t;\boldsymbol{x}_{2,0})$ of the system (8), starting at different initial conditions $\boldsymbol{x}_{1,0} \neq \boldsymbol{x}_{2,0}$, are called indistinguishable when

$$\boldsymbol{h}(\boldsymbol{x}_1(t)) = \boldsymbol{h}(\boldsymbol{x}_2(t)), \quad t \geq t_0 \tag{10}$$

holds true, where $\boldsymbol{h}(\cdot)$ denotes the output map. If (10) implies that $\boldsymbol{x}_1(t) = \boldsymbol{x}_2(t), \forall t \geq t_0$, then the system is called completely observable. Further, the system is called detectable when for indistinguishable trajectories it holds that

$$\lim_{t\to\infty} \boldsymbol{\epsilon}(t) = 0 \quad \text{where} \quad \boldsymbol{\epsilon}(t) = \boldsymbol{x}_2(t) - \boldsymbol{x}_1(t), \tag{11}$$

meaning that the indistinguishable trajectories converge to each other asymptotically.

Obviously, detectability is a weaker property than observability. Nevertheless, analyzing detectability can provide more model insight and can also be utilized in the observer design (see, e.g. (Moreno and Dochain, 2008; Jerono et al., 2021b)). Note that in the context of this observability analysis the term completely will always refer to $\mathbb{R}_{\geq 0}$ because the cell distributions are described by density functions and the biomass and substrate components are concentrations.

The analysis of the indistinguishable trajectories of the mass balance model (6) and the cell population balance model (1) and (1c) is carried out for biomass measurements

$$h(\boldsymbol{x}) = b = \Delta_m \sum_{i}^{z} m_i n_i = \boldsymbol{c}^T \boldsymbol{n} \tag{12}$$

where $\boldsymbol{c}^T = [\Delta_m m_1, ..., \Delta_m m_z]$. This is motivated by the fact that biomass is an available measurement in most bioreactors thanks to optical density sensors.

4.1 Indistinguishable trajectories of the mass balance model

Although the indistinguishable trajectories of the mass balance model (6) have been investigated in the literature (see, e.g (Schaum et al., 2005)) it is recalled here and serves as an important preliminary for further analysis.

Proposition 1. The mass balance model (6) with biomass measurements ($h(\boldsymbol{x}_b) = b$) and $\boldsymbol{x}_b = [b,\, s]^T$ is completely observable in $\mathbb{R}^2_{>0}$.

Proof. The error dynamics of indistinguishable trajectories $\dot{\boldsymbol{\epsilon}} = [\dot{\epsilon}_b,\, \dot{\epsilon}_s]^T$ of the mass balance model (6) with $\boldsymbol{x}_b = [b,\, s]^T$ read

$$\dot{\epsilon}_b = \rho(s+\epsilon_s)(b+\epsilon_b) - \rho(s)b \tag{13a}$$
$$\dot{\epsilon}_s = -\rho(s+\epsilon_s)(b+\epsilon_b) + \rho(s)b \tag{13b}$$
$$0 = h(\boldsymbol{x}_b + \boldsymbol{\epsilon}) - h(\boldsymbol{x}_b) = \epsilon_b. \tag{13c}$$

From the measurement equation (13c) one can directly conclude that the indistinguishable trajectories have to fulfill $\epsilon_b = 0$. Thus (13a) reduces to

$$\dot{\epsilon}_b = [\rho(s+\epsilon_s) - \rho(s)]b.$$

Note that for $\epsilon_b = 0, \forall t \geq t_0$ also $\dot{\epsilon}_b = 0$ has to hold. Due to the fact that $\rho(s)$ is given by the strictly increasing function (2) and considering $b \neq 0$ in a bioreactor, only $\boldsymbol{\epsilon} = [\epsilon_b,\, \epsilon_s]^T = \boldsymbol{0}^T$ can lead to the same system output. Therefore the mass balance model is completely observable within the state space $\boldsymbol{x}_b \in \mathbb{R}^2_{>0}$. □

Having this result, the substrate concentration s of the model (1a) and (1c) can be estimated by an observer operating on biomass measurements regardless of the cell population dynamics.

4.2 Indistinguishable trajectories of the discretized cell population balance model

The error dynamics of indistinguishable trajectories $\dot{\boldsymbol{\epsilon}} = [\dot{\boldsymbol{\epsilon}}_n,\, \dot{\epsilon}_s]^T$ of the population model (8) are given by

$$\dot{\boldsymbol{\epsilon}}_n = \boldsymbol{A}(s+\epsilon_s)(\boldsymbol{n}+\boldsymbol{\epsilon}_n) - \boldsymbol{A}(s)\boldsymbol{n} \tag{14a}$$

$$\dot{\epsilon}_s = -\rho(s+\epsilon_s)\Delta_m \sum_{i=1}^{z} m_i(n_i + \epsilon_{n,i}) + \rho(s)\Delta_m \sum_{i=1}^{z} m_i n_i \tag{14b}$$

$$0 = h(\boldsymbol{x}+\boldsymbol{\epsilon}) - h(\boldsymbol{x}) = \Delta_m \sum_{i=1}^{z} m_i \epsilon_{n,i} = \boldsymbol{c}^T \boldsymbol{\epsilon}_n = \epsilon_b. \tag{14c}$$

Clearly, at this point, $\boldsymbol{\epsilon}_n = 0$ cannot be concluded from the measurement equation (14c). Nevertheless, the analysis of indistinguishable trajectories of the mass balance model from the previous subsection implies that from $\dot{\epsilon}_b = 0$ it follows that $\epsilon_s = 0$ (Proposition 1) so that (14a) reduces to

$$\dot{\boldsymbol{\epsilon}}_n = \boldsymbol{A}(s)\boldsymbol{\epsilon}_n = \rho(s)\boldsymbol{\mathcal{A}}\boldsymbol{\epsilon}_n. \tag{15}$$

Furthermore one has that

$$\begin{aligned} \epsilon_b &= 0 = \boldsymbol{c}^T \boldsymbol{\epsilon}_n \\ \dot{\epsilon}_b &= 0 = \boldsymbol{c}^T \boldsymbol{A}(s)\boldsymbol{\epsilon}_n. \end{aligned} \tag{16}$$

Note that (16) also requires that $\frac{\mathrm{d}^{i-1}}{\mathrm{d}t^{i-1}}\epsilon_b = 0$ for $i = 1, ..., z$. This leads to the set of algebraic equations

$$\begin{aligned} \epsilon_b &= 0 = \boldsymbol{c}^T \boldsymbol{\epsilon}_n \\ \dot{\epsilon}_b &= 0 = \boldsymbol{c}^T \boldsymbol{A}(s)\boldsymbol{\epsilon}_n \\ &\vdots \\ \frac{\mathrm{d}^{z-1}}{\mathrm{d}t^{z-1}}\epsilon_b &= 0 = \frac{\mathrm{d}^{z-2}}{\mathrm{d}t^{z-2}}\boldsymbol{c}^T \boldsymbol{A}(s)\boldsymbol{\epsilon}_n. \end{aligned} \tag{17}$$

Observing that the dynamics (7a) are linear in $\boldsymbol{n}$ motivates to analyze (17) first for the case of constant growth rates $\frac{\mathrm{d}}{\mathrm{d}t}\rho(s) \approx 0$, i.e. $s \gg K_s$ (see (2)). Hence, the set of algebraic equations given by (17) and (14c) can be written as

$$\boldsymbol{0} = \begin{bmatrix} \boldsymbol{c}^T \\ \vdots \\ \boldsymbol{c}^T \boldsymbol{A}(s)^{z-1} \end{bmatrix} \boldsymbol{\epsilon}_n = \mathcal{O}_{z\times z}(s)\boldsymbol{\epsilon}_n \tag{18}$$

where indeed $\mathcal{O}_{z\times z}(s)$ has the structure of the classical Kalman observability matrix for linear systems. Obviously for $\rho(s) = 0$ equation (18) is fulfilled for arbitrary $\boldsymbol{\epsilon}_n$ (see (9)). Considering $\rho(s) > 0$ the question remains if the matrix $\mathcal{O}_{z\times z}(s)$ has full rank $\forall t \geq t_0$ in order to prove $\boldsymbol{\epsilon}_n = \boldsymbol{0}$ when (10) holds. From a structural point of view $\mathcal{O}_{z\times z}(s)$ has full rank, i.e. there exists at least one parameter-substrate combination such that $\mathrm{rank}(\mathcal{O}_{z\times z}(s)) = z$ due to the upper triangular structure of $\boldsymbol{A}(s)$ with additional elements on the first lower main diagonal originating from the first-order upwind finite differences. This also coincides with the structural observability property for the class of considered systems (see, e.g. (Jerono et al., 2021a)). Nevertheless, having $\frac{\mathrm{d}}{\mathrm{d}t}\rho(s) \neq 0$ (17) reads by utilizing the higher order product rule

$$\begin{aligned} \epsilon_b &= 0 = \boldsymbol{c}^T \boldsymbol{\epsilon}_n \\ \dot{\epsilon}_b &= 0 = \boldsymbol{c}^T \boldsymbol{A}(s)\boldsymbol{\epsilon}_n \\ &\vdots \\ \frac{\mathrm{d}^{z-1}}{\mathrm{d}t^{z-1}}\epsilon_b &= 0 = \boldsymbol{c}^T \sum_{k=0}^{z-2} \binom{z-2}{k} \left[\frac{\mathrm{d}^k}{\mathrm{d}t^k}\boldsymbol{A}(s)\right] \left[\frac{\mathrm{d}^{z-2-k}}{\mathrm{d}t^{z-2-k}}\boldsymbol{\epsilon}_n\right] \end{aligned} \tag{19}$$

where $\binom{z-2}{k}$ is the binomial coefficient. Again, structurally speaking, the set of algebraic equations (19) has the unique solution $\boldsymbol{\epsilon}_n = \mathbf{0}$. From an analytical point of view the problem is addressed in the next section for the special case of equal partitioning of the cells in the cell population model.

5. SPECIAL CASE OF EQUAL PARTITIONING

In order to analyze (19) it is assumed that cells can only divide into cells of equal size. As shown in (Diekmann et al., 1984) the model equation (1a) for this special case are given by

$$\begin{aligned}\partial_t n(m,t) &= -\partial_m[r(m,s)n(m,t)] - \Gamma(m,s)n(m,t) \\ &\quad + 4\Gamma_e(2m,s)n(2m,t) \end{aligned} \tag{20a}$$

$$n(m^*,t) = 0, \quad n(m,0) = n_0(m) \tag{20b}$$

having the structure of a classical convection reaction system where

$$\Gamma_e(2m,s) = \begin{cases}\Gamma(2m,s), & \text{if} \quad 2m \in [m_*,\, m^*] \\ 0, & \text{else.}\end{cases} \tag{21}$$

The factor 4 in equation (20a) originates from requiring mass conservation of cell division and birth as also discussed in (Diekmann et al., 1984). Nevertheless, applying the same discretization scheme as in Section 3 one will end up with the same structure as in equation (8), but the matrix will have more elements equal to zero in the upper triangular part due to the restriction of equal partitioning. One problem with equal partitioning for the discretized model equations is that mother cells being on an odd discretization point of the grid do not have daughter cells on the discretization grid. Therefore, it is further assumed that only mother cells with an even number on the discretization grid can divide into daughter cells. This leads to the dynamics

$$\begin{aligned}\dot{n}_i &= -\frac{1}{\Delta_m}\rho(s)(m_i n_i - m_{i-1}n_{i-1}) + \\ &\quad - \bar{\Gamma}(m_i,s)n_i + 2\bar{\Gamma}(m_{2i},s)n_{2i} \\ n_{z+1} &= 0, \quad n_i(0) = n_{i,0}, \quad s(0) = s_0\end{aligned} \tag{22}$$

where $\bar{\Gamma}(m_i,s)$ is given by

$$\bar{\Gamma}(m_i,s) = \begin{cases}\Gamma(m_i,s), & \text{if} \quad i \in 2\mathbb{N} \\ 0, & \text{if} \quad i \in 2\mathbb{N}+1.\end{cases} \tag{23}$$

Following this assumption the factor of 4 in (20a) becomes 2 again in the discretized case, because of (23). Equation (22) can again be rewritten in terms of

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix}\boldsymbol{A}_z(s)\boldsymbol{n} \\ f_s(\boldsymbol{x})\end{bmatrix}, \; \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^{z+1} \tag{24}$$

where $\boldsymbol{A}_z(s)$ now has the structure

$$\boldsymbol{A}_z(s) = \begin{bmatrix} -g_1^* & 2\gamma_2^* & 0 & 0 & 0 & 0 & \cdots & 0 \\ g_1^* & -g_2^* - \gamma_2^* & 0 & 2\gamma_4^* & 0 & 0 & \cdots & 0 \\ 0 & g_2^* & -g_3^* & 0 & 0 & 2\gamma_6^* & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

with $g_i^* = \frac{1}{\Delta_m}\rho(s)m_i$ and $\gamma_i^* = \bar{\Gamma}(m_i,s) = \gamma_i\rho(s)m_i$. The physiological properties of the cell cycle are still captured by these model dynamics. In fact the upper triangular part of $\boldsymbol{A}_z(s)$ has more elements equal to zero compared to $\boldsymbol{A}(s)$ in (9), which represents a worse case from a structural point of view. In order to address the dynamics of indistinguishable trajectories, two cases of interior discretization points $z = 2$ and $z = 4$ are considered. These cases refer to a rough classification of cells depending on their mass, which is meaningful when product accumulation properties of cells are related to specific mass intervals. For $z = 2$ the evaluation of (18) is handy. Note that for this case, on an equally spaced grid, the cell population model (7) is consistent with the equal partitioning assumption by default. Additionally, only the first time derivative of the system output has to be calculated, so that in this particular case (19) coincides with (17), i.e the problem can be addressed by evaluating the rank condition of $\mathcal{O}_{z\times z}(s)$ in (18). The $z = 4$ case leads to a more general consideration of cell division where the structure of $\boldsymbol{A}_z(s)$ in the case of equal partitioning is already taken into account.

5.1 Case $z = 2$

The system matrix reads

$$\boldsymbol{A}_2(s) = \begin{bmatrix} -g_1^* & 2\gamma_2^* \\ g_1^* & -\gamma_2^* \end{bmatrix}. \tag{25}$$

Note that the bottom right entry of $\boldsymbol{A}_2(s)$ does not contain the growth term. This is because the cells can only divide at this point in order to satisfy the boundary condition $n_{z+1} = 0$.

Proposition 2. The cell population model (22) with $z = 2$ and biomass measurements ($b = h(\boldsymbol{x})$) is completely observable in $\mathbb{R}^2_{>0}$.

Proof. In the considered case $\mathcal{O}_{2\times 2}(s)$ from equation (18) evaluates to

$$\begin{bmatrix} \Delta_m m_1 & \Delta_m m_2 \\ \rho(s)(m_1 m_2 - m_1^2) & \rho(s)\Delta_m(2\gamma_2 m_1 m_2 - m_2^2\gamma_2) \end{bmatrix}.$$

Having an equally spaced grid and by the assumption of equal partitioning $m_2 = 2m_1$ holds true. Thus the matrix can be written as

$$\mathcal{O}_{2\times 2}(s) = \begin{bmatrix} \Delta_m m_1 & \Delta_m 2m_1 \\ \rho(s)m_1^2 & 0 \end{bmatrix}. \tag{26}$$

The determinant reads

$$\det(\mathcal{O}_{2\times 2}(s)) = -2\Delta_m m_1^3 \rho(s), \tag{27}$$

so that $\mathcal{O}_{2\times 2}(s)$ is of full rank as long as $\rho(s) > 0$ holds true. □

Note that this result also holds for the cell population model (7a) due to the equivalence for $z = 2$.

5.2 Case $z = 4$

The system matrix reads

$$\boldsymbol{A}_4(s) = \begin{bmatrix} -g_1^* & 2\gamma_2^* & 0 & 0 \\ g_1^* & -g_2^* - \gamma_2^* & 0 & 2\gamma_4^* \\ 0 & g_2^* & -g_3^* & 0 \\ 0 & 0 & g_3^* & -\gamma_4^* \end{bmatrix}. \tag{28}$$

Evaluating (19) for the considered case leads to

$$
\begin{aligned}
\epsilon_b = 0 &= \boldsymbol{c}^T \boldsymbol{\epsilon}_n \\
\dot{\epsilon}_b = 0 &= \boldsymbol{c}^T \boldsymbol{A}_4(s) \boldsymbol{\epsilon}_n \\
\frac{\mathrm{d}^2}{\mathrm{d}t^2}\epsilon_b = 0 &= \boldsymbol{c}^T \left[\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}_4(s)\right] \boldsymbol{\epsilon}_n + \boldsymbol{c}^T \boldsymbol{A}_4(s) \dot{\boldsymbol{\epsilon}}_n \\
\frac{\mathrm{d}^3}{\mathrm{d}t^3}\epsilon_b = 0 &= \boldsymbol{c}^T \left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{A}_4(s)\right] \boldsymbol{\epsilon}_n \\
&\quad + 2\boldsymbol{c}^T \left[\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}_4(s)\right] \dot{\boldsymbol{\epsilon}}_n + \boldsymbol{c}^T \boldsymbol{A}_4(s) \left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{\epsilon}_n\right].
\end{aligned} \tag{29}
$$

Note that the time derivatives of the growth rate $\rho(s)$ appearing in $\boldsymbol{A}(s)$ can be calculated by (6b) and (2), but at this point, the further evaluation becomes more involved. Nevertheless, one can see that the set of equations (29) remains linear in $\boldsymbol{\epsilon}_n$ even for higher order derivatives. Therefore, a rank condition can be evaluated (for example with some suitable software) to check if for the indistinguishable trajectories $\boldsymbol{\epsilon}_n = \boldsymbol{0}$ holds true.

Here, for simplification purposes, the case of constant growth rates is assumed again.

Proposition 3. The cell population model (22) with $z = 4$ and biomass measurements ($b = h(\boldsymbol{x})$) is completely observable in $\mathbb{R}^4_{>0}$ for constant growth rates $\rho(s)$ and $s \neq 0$.

Proof. Utilizing $m_1 = \frac{1}{i} m_i$, $i \in \{2, 3, 4\}$ the matrix $\mathcal{O}_{4\times4}(s)$ from (18) evaluates to

$$
\mathcal{O}_{4\times4}(s) = \begin{bmatrix}
\Delta_m m_1 & \Delta_m 2m_1 & \Delta_m 3m_1 & \Delta_m 4m_1 \\
m_1^2\rho(s) & 2m_1^2\rho(s) & 3m_1^2\rho(s) & 0 \\
\dfrac{m_1^3\rho(s)^2}{\Delta_m} & \dfrac{2m_1^3\rho(s)^2}{\Delta_m} & -\dfrac{9m_1^3\rho(s)^2}{\Delta_m} & \dfrac{16\gamma_4 m_1^3\rho(s)^2}{\Delta_m} \\
\dfrac{m_1^4\rho(s)^3}{\Delta_m^2} & -\dfrac{22m_1^4\rho(s)^3}{\Delta_m^2} & \mathcal{O}_{3,4} & \mathcal{O}_{4,4}
\end{bmatrix}
$$

where

$$
\mathcal{O}_{3,4} = \frac{(48\gamma_4\Delta_m + 27)m_1^4\rho(s)^3}{\Delta_m^2}
$$

$$
\mathcal{O}_{4,4} = -\frac{(64\gamma_4^2\Delta_m - 16\gamma_4)m_1^4\rho(s)^3}{\Delta_m}.
$$

The determinant is given by

$$
\det(\mathcal{O}_{4\times4}(s)) = 1152\frac{m_1^{10}\rho(s)^6}{\Delta_m^2}, \tag{30}
$$

revealing that $\mathcal{O}_{4\times4}(s)$ has also full rank as long $\rho(s) > 0$ holds true. Therefore $\epsilon_b = 0$, $\forall t \geq t_0$ requires $\boldsymbol{\epsilon}_n = \boldsymbol{0}$, $\forall t \geq t_0$ such that there exist no indistinguishable trajectories in the time interval $t \in [t_0, \infty)$. □

Comparing (27) with (30) it looks convenient to elaborate a general formula of the determinant depending on z, but this will not be further addressed here.

For the considered number of interior discretization points the complete observability within the state space $\boldsymbol{x} \in \mathbb{R}^z_{>0}$ can be proven by contradiction of the existence of indistinguishable trajectories for $s \neq 0$. In the case of $z = 4$ the carried out analysis is restricted to constant growth rates, i.e. $s \gg K_s$. It has to be pointed out that in both cases the observability property is lost for $\rho(s) = 0$. From a practical point of view this means that in a batch experiment the observer has to converge before the substrate is completely consumed. Additionally, an observer which is designed based on the observability map, like the one presented in the next section, might run into numerical issues due to an ill condition when $\rho(s)$ approaches zero. Therefore, it is appropriate to design a stopping criteria for the observer gain calculation.

6. SIMULATION RESULTS

Based on the previous analysis of the indistinguishable trajectories an observer is designed to estimate the cell distribution density operating on biomass measurements for the model equations (22) when $z = 2$ is considered. Note that the analysis in section 4 leads to an evaluation of a rank condition due to the linearity of the system with respect to $\boldsymbol{n}$. In fact, the system is nonlinear due to the coupling to the nonlinear substrate dynamics. Given that the substrate concentration can be estimated based on a mass balance observer operating on biomass measurements, a reduced observer structure can be considered to estimate the cell population, i.e. there will be no correction term based on the substrate concentration allowing the interpretation as a (known) system parameter. Nevertheless, since the substrate concentration is still a function of time, the resulting observer design problem for the cell population model will be linear time–variant. One possibility of designing an observer for this class of systems is given by utilizing Ackermann formula for linear time–variant systems (Silverman and Meadows, 1967; Freund, 2013) which is based on the (time–variant) observability map of the system given by

$$
\mathcal{O}_t = \begin{bmatrix} M_A^0 \boldsymbol{c}^T \\ \vdots \\ M_A^{z-1} \boldsymbol{c}^T \end{bmatrix} \tag{31}
$$

with the operator $M_A^0 \boldsymbol{c}^T = \boldsymbol{c}^T$, $M_A^1 \boldsymbol{c}^T = \dot{\boldsymbol{c}}^T + \boldsymbol{c}^T \boldsymbol{A}$, $M_A^k \boldsymbol{c}^T = M_A(M_A^{k-1}\boldsymbol{c}^T)$ and the linear, possibly time-varying, measurement equation $\boldsymbol{c}^T$. The observer dynamics are given by

$$
\begin{aligned}
\dot{\hat{\boldsymbol{x}}} &= \boldsymbol{A}_2(s)\hat{\boldsymbol{x}} + \boldsymbol{l}(t)(y - \hat{y}) \\
\hat{y} &= \boldsymbol{c}^T \hat{\boldsymbol{x}}
\end{aligned} \tag{32}
$$

with $\hat{\boldsymbol{x}} = [\hat{n}_1, \hat{n}_2]^T \in \mathbb{R}^2$. The correction gain $\boldsymbol{l}(t)$ reads

$$
\boldsymbol{l}(t) = \frac{1}{\bar{c}(t)}[p_0 N_A^0 + ... + p_{z-1} N_A^{z-1} + N_A^z] \circ \boldsymbol{v}(t) \tag{33}
$$

$$
\boldsymbol{v}(t) = \mathcal{O}_t^{-1} \begin{bmatrix} 0 & ... & \bar{c}(t) \end{bmatrix}^T \tag{34}
$$

with the operator $N_A^0 \boldsymbol{v} = \boldsymbol{v}$, $N_A^1 \boldsymbol{v} = -\dot{\boldsymbol{v}} + \boldsymbol{A}\boldsymbol{v}$, $N_A^k \boldsymbol{v} = N_A(N_A^{k-1}\boldsymbol{v})$, the coefficients p_j, $j \in \{0, ..., z-1\}$ of an Hurwitz polynomial and the design parameter $\bar{c}(t)$. The selected system and observer design parameters are listed in Table 1. The simulation results are presented in Figure 1. The initial states are $\boldsymbol{x} = [2, 0]^T$ for the real system and $\hat{\boldsymbol{x}} = [1.5, 0.5]^T$ for the observer. As

Table 1. Parameter list

Parameter	Value	Unit	Parameter	Value	Unit
k_s	0.02	h^{-1}	Δ_m	1	g
K_s	1	g/l	p_0	0.25	-
γ_2	20	–	p_1	1	-
m_1	1	g	ξ	10^{-3}	g/l
m_2	2	g	$\bar{c}$	$\det(O_t)$	$\mathrm{g}^4\mathrm{h}^{-1}$

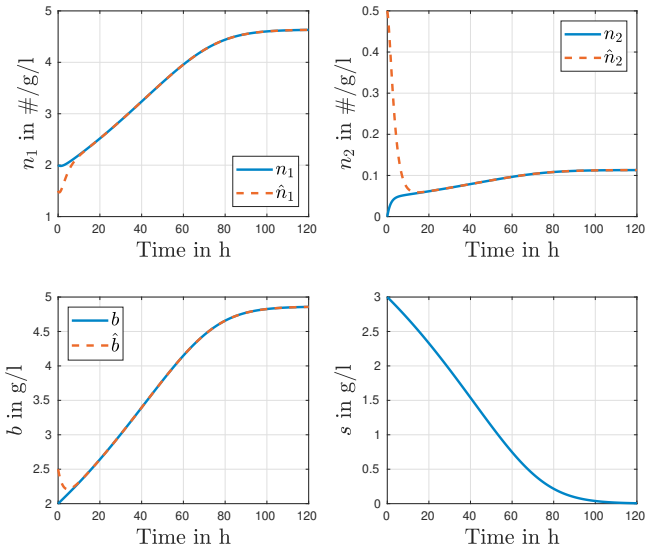

Fig. 1. Estimates (dashed–red) and real values (solid–blue) of the cell distribution density n_1, n_2, the biomass b and substrate concentration s in a batch scenario.

stated before, the observability property is lost when the substrate (and therefore $\rho(s)$) approaches zero. In order to avoid numerical issues in $\mathcal{O}_t^{-1}$, the calculation of $\boldsymbol{l}(t)$ is stopped and its values are set to zero when the substrate is below a certain threshold ξ. At this point the observer is only driven by the simulator, meaning that in order to provide reliable estimates, the observer has to converge before $s < \xi$ holds true.

7. CONCLUSION

The observability property of the finite-dimensional cell population balance equation with biomass measurements within the framework of the indistinguishable trajectories was addressed. The population balance equation is described by a partial integro-differential equation which is coupled with an ordinary differential equation for the substrate dynamics. For the semi-discretized model equations the dynamics of indistinguishable trajectories were derived. The dynamics were analyzed considering the special cases of constant growth rates and equal partitioning. It turned out that the analysis of indistinguishable trajectories can be re–cast into a rank condition of a matrix due to the linearity of the equations with respect to the cell distribution density. For the case of constant growth rates the analysis coincides with the classical Kalman observability condition for linear time–invariant systems. The analysis revealed that the observability property is guaranteed as long as the nutrient concentration is not depleted. Further, the observability property was established for a small number of discretization points. Based on the results an observer was designed for a bi–partitioned cell distribution and tested in simulations.

REFERENCES

Beniich, N., Abouzaid, B., and Dochain, D. (2018). On the existence and positivity of a mass structured cell population model. Applied Mathematical Sciences, 12(19), 921–934.

Diekmann, O., Heijmans, H.J., and Thieme, H.R. (1984). On the stability of the cell size distribution. Journal of Mathematical Biology, 19(2), 227–248.

Freund, E. (2013). Zeitvariable Mehrgrößensysteme, volume 57. Springer-Verlag.

Ibarra, S., Moreno, J., and Espinosa-Pérez, G. (2004). Global observability analysis of sensorless induction motor. Automatica, 40(6), 1079–1085.

Isidori, A., Sontag, E., and Thoma, M. (1995). Nonlinear control systems, volume 3. Springer.

Jerono, P., Schaum, A., and Meurer, T. (2021a). Moment-based kalman filter design for cell population balance models in batch fermentation processes. IFAC-PapersOnLine, 54(3), 19–24.

Jerono, P., Schaum, A., and Meurer, T. (2021b). Observability analysis and robust observer design for a continuous yeast culture. Journal of Process Control, 104, 62–73.

Liu, Y.Y., Slotine, J.J., and Barabási, A.L. (2013). Observability of complex systems. Proceedings of the National Academy of Sciences, 110(7), 2460–2465.

Mantzaris, N.V., Liou, J.J., Daoutidis, P., and Srienc, F. (1999). Numerical solution of a mass structured cell population balance model in an environment of changing substrate concentration. Journal of Biotechnology, 71(1-3), 157–174.

Mantzaris, N.V. and Daoutidis, P. (2004). Cell population balance modeling and control in continuous bioreactors. Journal of Process Control, 14(7), 775–784.

Moreno, J.A. and Dochain, D. (2005). Global observability and detectability analysis of uncertain reaction systems. IFAC Proceedings Volumes, 38(1), 37–42.

Moreno, J.A. and Dochain, D. (2008). Global observability and detectability analysis of uncertain reaction systems and observer design. International Journal of Control, 81(7), 1062–1070.

Moreno, J.A., Rocha-Cózatl, E., and Wouwer, A.V. (2014). A dynamical interpretation of strong observability and detectability concepts for nonlinear systems with unknown inputs: application to biochemical processes. Bioprocess and biosystems engineering, 37, 37–49.

Nijmeijer, H. and Van der Schaft, A. (1990). Nonlinear dynamical control systems, volume 175. Springer.

Schaum, A. and Jerono, P. (2019). Observability analysis and observer design for a class of cell population balance models. IFAC-PapersOnLine, 52(2), 189–194.

Schaum, A. and Moreno, J.A. (2007). Dynamical analysis of global observability properties for a class of biological reactors. IFAC Proceedings Volumes, 40(4), 213–218.

Schaum, A., Moreno, J.A., and Vargas, A. (2005). Global observability and detectability analysis for a class of nonlinear models of biological processes with bad inputs. In 2005 2nd International Conference on Electrical and Electronics Engineering, 343–346. IEEE.

Silverman, L.M. and Meadows, H. (1967). Controllability and observability in time-variable linear systems. SIAM Journal on Control, 5(1), 64–73.

Tsuchiya, H.M., Fredrickson, A.G., and Aris, R. (1966). Dynamics of microbial cell populations. Advances in Chemical Engineering, 6(C), 125–206.

Zeitz, M. (1990). Canonical forms for nonlinear systems. In Nonlinear Control Systems Design 1989, 33–38. Elsevier.