Set Membership Estimation for Dynamic Flux Balance Models

Authors:

Xin Shen, Hector Budman

Date Submitted: 2021-10-21

Keywords: set membership estimation, dynamic flux balance model, multiparametric programming, multiplicity, weighted primal dual method, variable structure system

Abstract:

A set membership estimator (SME) based on limited number of measurements is proposed for estimating metabolite concentrations using dynamic flux balance models (DFBMs). To deal with multiplicity of solutions of the DFBM, a weighted primal dual method is used to find solutions that best fit the data. Multiparametric nonlinear programming is applied to propagate uncertainty in initial concentrations along a batch/fed-batch operation. The proposed method has been applied to E. coli batch and fed-batch fermentation without noise.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):	LAPSE:2021.0786
Citation (this specific file, latest version):	LAPSE:2021.0786-1
Citation (this specific file, this version):	LAPSE:2021.0786-1v1

CSChE Systems & Control Transactions

Open Access | Original Research

Conference Proceedings
Canadian Chemical Engineering Conference 2021
October 24-27, Montréal, Québec, Canada

PSEcommunity.org/LAPSE:2021.0786

PSE PRESS

Set Membership Estimation for Dynamic Flux Balance Models

Xin Shen and Hector Budman*

Department of Chemical Engineering, University of Waterloo, Waterloo, ON, Canada
* Corresponding Author: hbudman@uwaterloo.ca.

ABSTRACT

A set membership estimator (SME) based on limited number of measurements is proposed for estimating metabolite concentrations using dynamic flux balance models (DFBMs). To deal with multiplicity of solutions of the DFBM, a weighted primal dual method is used to find solutions that best fit the data. Multiparametric nonlinear programming is applied to propagate uncertainty in initial concentrations along a batch/fed-batch operation. The proposed method has been applied to E. coli batch and fed-batch fermentation without noise.

Keywords: set membership estimation, dynamic flux balance model, multiparametric programming, multiplicity, weighted primal dual method, variable structure system

Date Record: Original manuscript received October 13, 2021. Published October 21, 2021

1. INTRODUCTION

Dynamic flux balance models (DFBMs) are one type of genome-based structural models that has gained attention in recent years due to their ability to describe the bioprocess with fewer tuning parameters in comparison with other modelling techniques. In DFBMs, cells are regarded as optimizing agents able to distribute resources to maximize/minimize a biological objective. Accordingly, DFBMs models can be formulated as a linear programing (LP) constrained optimization problem which solution serves as input to mass balances of metabolites.

However, due to multiplicity of solutions of the LP, the possible trajectories of the state variables, i.e. metabolites concentrations, for given initial condition are not unique for DFBMs. In a previous investigation by the authors, a weighted primal dual method (WPDM) was used to solve the multiplicity problem by introducing the weights that drive the optimization search to solutions that best fit the experimental data [1]. The WPDM approximates the LP by nonlinear convex programming. To propagate uncertainty in initial bounds on concentrations during the batch multiparametric nonlinear programming is applied to convert the system into a variable structure one.

A set membership estimation approach is used based on few available measurements to estimate bounds for all states. The proposed methods have been applied to fed-batch and batch fermentation of *E. coli* model.

2. METHODS

2.1 Dynamic Flux Balance Models.

The governing equation of DFBMs can be defined as Eq. (1) – Eq. (3) [2].

$$x_{k+1} = B(u_k, x_{vol,k})x_k + \Delta t x_{bio,k} S v_k + h(u_k, x_{vol,k}) \quad (1)$$

$$y_k = [x_{bio,k}\ x_{vol,k}]^T \quad (2)$$

$$x_0 \in \mathcal{P}_0 \quad (3)$$

where x_k is state vector of n_x state variables at time step k, including state biomass concentration $x_{bio,k}$ and volume $x_{vol,k}$. Δt is discrete time step size. S is stoichiometry matrix of metabolic reactions. v_k is metabolic flux. B and h are matrix -valued and vector-valued functions of input variable u_k and volume. In this research, we assume biomass concentration and volume can be measured. The initial states are uncertain but are known to lie within the set $\mathcal{P}_0$.

The flux v_k is obtained by solving LP in Eq. (4):

$$\min_{v_k} c^T v_k$$
$$\text{Subject to } Av_k \leq F\theta_k(x_k, y_k, u_k) + b \quad (4)$$

where matrix $A \in \mathbb{R}^{m \times n_x}$ and $F \in \mathbb{R}^{m \times n_\theta}$ are constant matrix and b is constant vector. c is objective coefficient vector. $\theta_k = H(y_k, u_k)x_k + g(y_k, u_k)$. $H(y_k, u_k)$ and $g(y_k, u_k)$ are matrix -valued and vector-valued functions of input variable u_k and measurements.

2.2 Weighted Primal Dual Method

Multiplicity of solutions is frequently observed in DFBM in view of insufficient number of constraints. Since the solution of LP is usually not unique for Eq. (4), different trajectories can be obtained by different LP solvers. To deal with the multiplicity issue, a weighted primal dual method (WPDM) was proposed in a previous work by the authors [1]. The

WPDM involves the addition of a weighted barrier function into the optimization cost as per Eq. (5), where w are interior-point weights and z are slack variables. μ is set as a very small number, $\mu = 1e^{-8}$. For simplicity, subscript k is omitted. Since the WPDM has strictly convex form objective, the solution is unique. When multiple solutions coexist in the original LP defined in Eq. (4), different unique solutions can be obtained corresponding to different weights. The weights in WPDM are calibrated to fit available experimental data.

$$\begin{aligned} &\min_{v} c^T v - \mu \sum_{i=1}^{m} w_i \ln(z_i) \\ &\text{Subject to } Av + z = F\theta + b \\ &z > 0 \end{aligned} \tag{5}$$

2.3 Multiparametric Programming

2.3.1 Multiparametric Nonlinear Programming

Since θ in Eq. (5) varies with time step k, θ can be regarded as varying parameters. Accordingly, Eq. (5) can be treated as a multiparametric nonlinear strictly convex programming problem defined as follows:

$$\begin{aligned} &\min_{v} f(v, \theta) \\ &\text{Subject to } Gv < K\theta + d \\ &G_e v = K_e \theta + d_e \\ &\theta \in \Theta \end{aligned} \tag{6}$$

Assume the Θ can be divided into n_Θ nonoverlapping sets as Θ^1, ..., Θ^{n_Θ}. These sets are defined as critical regions. For given θ_0 in a critical region Θ^j, the corresponding optimal solution is v_0. So at θ_0 and v_0, the solution can be approximated by a local quadratic programming problem as Eq. (7) [3].

$$\begin{aligned} &\min_{v} \frac{1}{2}(v - v_0)^T Q(v - v_0) + (L^T + (\theta - \theta_0)^T N)(v - v_0) \\ &\quad + r(v_0, \theta_0) \\ &\text{Subject to } Gv < K\theta + d \\ &G_e v = K_e \theta + d_e \\ &\theta \in \Theta^j \end{aligned} \tag{7}$$

where $H = \nabla^2_{vv} f(v_0, \theta_0)$, $L = \nabla_v f(v_0, \theta_0)$, $N = \nabla^2_{\theta v} f(v_0, \theta_0)$, $r(v_0, \theta_0) = f(v_0, \theta_0) + \nabla^T_\theta f(v_0, \theta_0)(v - v_0) + \frac{1}{2}(\theta - \theta_0)\nabla^2_{\theta\theta} f(v_0, \theta_0)(\theta - \theta_0)$. The optimal solution of Eq. (7) has affine form as $v^j = A^j_Q \theta + b^j_Q$, where superscript j denotes that for each critical regions a different analytical expression can be used to calculate the optimum. Different published multiparametric nonlinear programming codes can be used to obtain the constant matrix A_Q and b_Q. The critical regions can be further divided into smaller critical regions to improve the solution accuracy.

2.3.2 Variable Structure System

The state equation in Eq. (1) can be rewritten as Eq. (8) and Eq. (9).

$$\begin{aligned} x_{k+1} &= B(u_k, x_{vol,k})x_k + \Delta t x_{bio,k} S(A^j_Q \theta_k + b^j_Q) \\ &\quad + h(u_k, x_{vol,k}) \end{aligned} \tag{8}$$

$$\theta_k(x_k, y_k, u_k) \in \Theta^j \tag{9}$$

Since θ_k is a function of x, different critical region corresponds to different state space models. Therefore, the state equation is different in different positions of the state space thus defining a piecewise continuous system representation that is referred to as a variable structure system [4].

2.3 Set Membership Estimation (SME)

For the variable structure system defined above set membership estimation can be used to propagate the initial value uncertainty $\mathcal{P}_0$ to estimate bounds on concentrations along the batch based on few measurements. Set membership estimation is based on set operations, usually including linear mapping, projection, translation, Minkowski addition, intersection, union, and outer approximation. Multi-Parametric Toolbox 3 [5] and bensolve tools [6] are used to preform the set operations. By taking the measurements into consideration, post estimate sets of states at different time steps can be obtained. Through projections of the obtained sets, the upper bounds and lower bounds of states can be determined.

3. RESULTS AND DISCUSSION.

The proposed set membership estimation has been applied to an E. coli fermentation example for both batch and fed-batch operations. To test the multiplicity issue, an artificial constraint is added into the DFBM of *E. coli* on purpose. Figure 1 presents the results of simulation of batch fermentation. Assume only volume and biomass concentration are measured, the red dash lines are estimated bounds for unmeasured states which also cannot be estimated by classical observers. It was concluded that the proposed SME generates tight bounds that contain the true plant state denoted by black line. Similar results have been obtained for fed-batch operations.

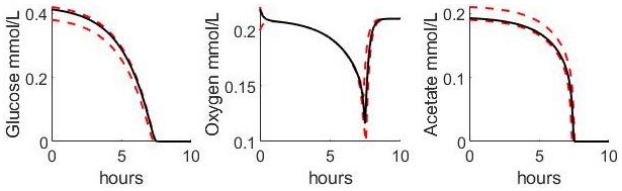

Figure 1. SME for batch fermentaiton of *E.coli*.

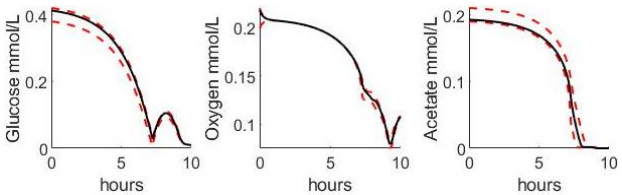

Figure 2. SME for fed-batch fermentaiton of *E.coli*..

4. CONCLUSIONS.

DFBMs based on WDPM solver can solve the multiplicity problem. Multiparametric nonlinear programming is used to convert the DFBMs based on WPDM solver into a variable structure system. Based on the simplified variable structure system, SME can propagate the initial value uncertainty by set operations and estimate the bounds of unmeasured states given the limited measurements.

ACKNOWLEDGEMENTS

We acknowledge the support from The Natural Sciences and Engineering Research Council of Canada (NSERC), Mitacs, and Sanofi Pasteur.

REFERENCES

1. Shen X, Budman H. A method for tackling primal multiplicity of solutions of dynamic flux balance models. *Comp Chem Eng*, 143:107070 (2020).
2. Mahadevan R, Edwards JS, Doyle FJ III. Dynamic flux balance analysis of diauxic growth in Escherichia coli. *Biophy J*, 83:3:1331-1340 (2002).
3. Johansen TA. On multi-parametric nonlinear programming and explicit nonlinear model predictive control. *Proceedings of the 41st IEEE Conference on Decision and Control*, 3:2768-2773 (2002).
4. Shen X, Budman H. A Type of Set Membership Estimation Designed for Dynamic Flux Balance Models. *Processes* 9:10:1762 (2021).
5. Herceg M, Kvasnica M, Jones CN, et al. Multi-parametric toolbox 3.0, *Euro Cont Conf*, 502-510 (2013).
6. Löhne A, Weißing B. Equivalence between polyhedral projection, multiple objective linear programming and vector linear programming. *Math Oper Res* 84:2:411-426 (2016).

This conference proceeding has not been peer reviewed.