Chemical Engineering Science 167 (2017) 18–28

Contents lists available at ScienceDirect

Chemical Engineering Science

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

Applications of Polynomial Chaos Expansions in optimization and control of bioreactors based on dynamic metabolic flux balance models ☆

Divya Kumar, Hector Budman *

Chemical Engineering Department, University of Waterloo, Waterloo, ON N2L3G1, Canada

HIGHLIGHTS

- A new model-based online robust control methodology maximizing the end point property of a fed-batch bioreactor is proposed.
- A new model-based offline robust optimization methodology for fed-batch bioreactor is proposed.
- Both the algorithms use Polynomials Chaos Expansions for uncertainty propagation and show superior performance than nominal counterparts.

ARTICLE INFO

Article history:
Received 8 September 2016
Received in revised form 24 January 2017
Accepted 13 March 2017
Available online 15 March 2017

Keywords:
Dynamic Flux Balance Model
Fed-batch bioreactor
Robust control
Robust optimization
Model predictive control
Polynomial Chaos Expansions

ABSTRACT

This work proposes model based approaches for on-line or off-line economic optimization of batch reactors in the presence of model error (uncertainty). Polynomial Chaos Expansions are used as an effective and computationally efficient tool to propagate the error in model parameters into the optimizations' cost functions. The computational efficiency of the proposed uncertainty propagation approach is essential for the implementation of the on-line approach that takes into account feedback corrections. The role of feedback, as applied in the on-line formulation, is proved to be instrumental for reducing conservatism of the optimization results. The proposed approaches can serve to design recipes for maximizing productivity in batch, fed-batch or perfusion operation of bioreactors.

1. Introduction

This work proposes model-based economic control and optimization approaches for bioreactor processes that are robust to model error. The key challenge in addressing robustness to model error is to propagate the uncertainty in model parameters onto the control or optimization objective. When using nonlinear dynamic first principles' models such propagation requires the use of Monte Carlo algorithms which are computationally demanding. To reduce the computational load we propose the use of Polynomial Chaos Expansions that permit quick calculation of the variance resulting from the process model mismatch.

Two different problems are tackled: i- on-line robust predictive control with an economic objective, referred also to as economic predictive control and ii- off-line robust optimization of an end point property.

Most of the reported studies on optimal operation of bioreactors involve offline model based optimization (Banga et al., 1997), without accounting for feedback corrections, (Frahm et al., 2002; Hjersted and Henson, 2006; Banga et al., 1997) or for robustness to model errors. The objective of these optimization strategies have been generally the maximization of a property at the end of the batch such as the productivity.

Traditionally, studies of optimization of bioreactors have used unstructured models that are based on simplistic substrate and biomass balances thus not accounting for detailed interactions between different nutrients. On the other hand, structured models that explicitly account for detailed interactions between nutrients and products, have gained increasing acceptance in the pharmaceutical industry motivating their use for control and optimization. For example, Dynamic Flux Balance Modeling (DFBM) has been applied successfully by Mahadevan et al. (2002), as an extension of MFA to describe the dynamic growth of *E. coli* on glucose and acetate. Hjersted and Henson (2006) used DFBM models representing the growth of *Saccharomyces Cerevisae* and Ethanol production on glucose for offline optimization of the fed-batch operation by implementing an optimal substrate feeding policy while reducing batch time or/and increasing

☆ Part of this work has been adapted from Kumar, D. & Budman, H. 2015.

* Corresponding author.
E-mail address: hbudman@uwaterloo.ca (H. Budman).

http://dx.doi.org/10.1016/j.ces.2017.03.035
0009-2509/

productivity. DFBM have been also used for biological process optimization, e.g. for the increase of productivity of microalgae (Flassig et al., 2016). A key advantage of DFBM models is that they require solving an LP problem with a relatively small number of rate limiting kinetic constraints as compared to other unstructured models that require the calibration of a larger number of kinetic expressions with many corresponding parameters. Thus, DFBMs are potentially less sensitive to experimental noise than other models but they can be sensitive to parametric uncertainty. Also, during fed-batch operation, a combination of factors such as non-ideal mixing or the presence of froth may contribute to additional model error and process disturbances. The importance of robustness for fed-batch bioreactor control has been stressed in Kuhlmann et al. (1998), due to (i) time varying behavior, (ii) un-modeled dynamics and (iii) large disturbances occurring in the process.

Due to the computational demand associated with Monte Carlo simulations for propagating uncertainty, approximations have been proposed. For example, conventional moment-equation (CME) approaches have been proposed where the stochastic function is approximated by power series of high order and then the statistical moments of the stochastic function can be formulated as deterministic equations based on these expansions. Nagy and Braatz (2007) have shown that Polynomial Chaos Expansions (PCE) is a computationally efficient alternative to Monte Carlo simulations for propagating uncertainty in dynamic models. PCE is based on orthogonal basis functions thus requiring smaller function evaluations for the calculation of numerical integrations needed for obtaining statistical moments as compared to the CME method mentioned above (Lu and Zhang, 2005). The computational advantages of PCEs for robust control and optimization (Nagy and Braatz, 2007; Kim et al., 2012; Kumar and Budman, 2014) derive from the availability of analytical formulae to compute the statistical moments (mean, variance, etc.) of variables described by such expansions.

In the current study, Polynomial Chaos Expansions is applied to propagate uncertainty onto a quality of interest for both an on-line robust optimal control and an off-line robust optimization problems. For both applications the process dynamics are modeled using DFBM and the parametric uncertainty is propagated using a PCE based approach. Since the DFBM model involves an LP, the resulting control strategy is obtained from the solution of a bi-level optimization problem involving the maximization of the economic objective subject to the LP solution. This bi-level optimization formulation poses challenges to the design of a robust strategy and the propagation of the uncertainty onto the solution; and a PCE based approach is proposed to address them. The proposed controller can be used in real-time application due to the low computational complexity resulting from the use of PCEs. Even when feedback is not used, fast uncertainty propagation is very important for the implementation of batch to batch optimization procedures where the time between the two batches available for performing optimization is limited.

The manuscript is organized as follows. Section 2 motivates the work by presenting the case study and provides background material on PCE. Section 3 and Section 4 present the on-line and off-line optimization approaches respectively followed by Conclusions in Section 5. Implementation of all presented algorithms are given in Kumar (2014).

2. Mathematical background

2.1. Motivation of the work: Case study

The motivation for this work is explained through the case study that was used involving a bioreactor where *E. coli* is grown on glucose with the objective of producing a particular protein. It will be assumed that the process can be operated in fed-batch or perfusion modes with a pre-defined total duration. Along the run nutrient is fed or perfusion of supernatant is used to control and/or optimize the operation. For simplicity it will be assumed that the protein of interest is proportional to the amount of biomass thus higher biomass at the end of the batch equates to higher productivity. Following this assumed correlation between biomass and product the explicit level of the latter will not be explicitly solved or discussed further in this example. In order to optimize the productivity of this process there are two main options available: i- to calculate on-line feeding and perfusion rates at each time interval that will maximize productivity at the end of the batch based on measurements that are available at regular intervals during the operation and ii- to calculate off-line recipes for feeding or perfusion rates' profiles along the operation that maximize productivity at the end of the batch and to apply these fixed profiles in all future batches. The first option will be referred heretofore as on-line optimization whereas the second option will be referred to as off-line control. The control problem treated here is also referred in the literature as economic control since the objective is to maximize a profit function (productivity) rather than to minimize errors with respect to a preselected set point as done in conventional control strategies.

A fundamental model referred to as dynamic metabolic flux model (DFBM) is used in this work for both on-line control and off-line optimization. DFBM is based on an a priori known network of m metabolites, $\boldsymbol{z_{mx1}}$, participating in n different reactions. Each reaction is associated to a flux, $\boldsymbol{v_{nx1}}$ given in units of mM of metabolite/hr/mM of cell. This network of reactions can be mathematically expressed in terms of a stoichiometric matrix ($\boldsymbol{A_{mXn}}$) for the corresponding vector of reaction fluxes ($\boldsymbol{v_{nx1}}$). The DFBM approach assumes that the cell acts as an agent that strives to optimally allocate available resources (nutrients) to maximize a given objective, e.g. the cellular growth rate μ. Other optimization objectives have also been reported, e.g. the redox potential, but this study considers only the cell growth. A simplified DFBM model developed by Mahadevan et al. (2002) for growth of *E. coli* on glucose is used. Fig. 1 shows the simplified metabolic network, with glucose (Glcxt), acetate (Ac) and oxygen (O_2) as the input and biomass (X) as the output.

It consists of 4 fluxes given by the vector $\boldsymbol{v}$ and 3 metabolites given by the vector $\boldsymbol{z}$ (Glcxt, Ac, O_2). The growth rate, μ as a function of the fluxes and the stoichiometric matrix $\mathbf{A_{mxn}}$, the stoichiometric matrix related to the 3 metabolites participating in the reactions leading to the biomass growth, are presented as follows:

$$\boldsymbol{A} = \begin{bmatrix} 0 & -9.46 & -9.84 & -19.23 \\ -35 & -12.92 & -12.73 & 0 \\ -39.43 & 0 & 1.24 & 12.12 \end{bmatrix} \tag{1}$$

and, $\mu = \sum_{i=1}^{4} v_i$

Assuming the cells try to continuously maximize growth, the fluxes v_i can be solved from the following optimization problem:

$$\max_{X, \boldsymbol{v_i}} \ \mu = \sum_{i=1}^{4} v_i \tag{2}$$

$$z_{i,} \geqslant 0, \ \forall i \in [1,3], \quad v_i \geqslant 0, \ \forall i \in [1,4]$$

$$|A^{Glcxt}\boldsymbol{v}| \leqslant \frac{GUR_{max} Z_{Glcxt}}{K_m + Z_{Glcxt} + \frac{Z_{Glcxt}^2}{K_I}} \frac{mmol}{gdw - hr}$$

$$-A^{O2}\boldsymbol{v} \leqslant OUR_{max},$$

$$A^{Ac}\boldsymbol{v} \leqslant 100$$

This biological system involves 3 distinct growth phases of *E. coli*, viz. (i) Aerobic growth on Glucose, (ii) Anaerobic growth