Model Predictive Control of bioreactors based on a reformulation of dynamic metabolic network models

Marius Fredriksena, Rafael David de Oliveiraa, Caroline Satye Nakamaa, Johannes Jäschkea

aDepartment of Chemical Engineering, Norwegian University of Science and Technology, NTNU, NO-7491 Trondheim, Norway

[johannes.jaschke@ntnu.no](mailto:johannes.jaschke@ntnu.no)

Abstract

The increasing popularity and utilization of dynamic Flux Balance Analysis (dFBA) models have allowed for the implementation of more advanced control structures, such as Model Predictive Control (MPC), for bioprocesses. The dFBA model is comprised of dynamic mass balance equations and an optimization that calculates the cell's metabolic fluxes. Thus, when MPC is applied, a bi-level optimization problem arises. Solving this bi-level optimization problem involves transforming the inner optimization into a set of algebraic equations using either the duality theory or the Karush–Kuhn–Tucker (KKT) optimality conditions. In this work, we assessed different reformulations of the dFBA model for continuous stirred tank (CSTR) bioreactors that would make the dFBA suitable for MPC applications. We conducted a case study of the *Escherichia coli* (*E. coli*) core metabolic network, where we applied MPC to the CSTR bioreactor. The controller was subjected to changes in setpoint and disturbances in the glucose feed concentration and the maximal glucose uptake for the cells. The MPC controller performed well, maintaining the biomass concentration close to the desired setpoint. Overall, the reformulation based on the penalized duality theory was more reliable than the penalized KKT reformulation.

**Keywords**: Bioprocess control, Dynamic Flux Balance Analysis, Model Predictive Control, Bi-level optimization

* 1. Introduction

Bioprocessing is an integral field of research as it is used in various food, chemical, and pharmaceutical industries (Doran, 2013). Bioprocessing refers to the use of cells and their components, like enzymes, to manufacture goods and destroy harmful waste. From its ancient origins in food production, bioprocessing has evolved to encompass the manufacturing of an extensive range of commercial products, from relatively cheap materials like industrial alcohol and organic solvents to more valuable products like antibiotics, vaccines, and therapeutic proteins (Doran, 2013).

The increasing competition, stricter regulations, and economic fluctuations in recent years have emphasized the importance of good process control and optimization, which is crucial for safe and efficient plant operations (Seborg et al., 2016). Current bioprocess control structures are recipe-based, with insufficient ability to handle uncertainties (Jabarivelisdeh et al., 2020). However, recent improvements in genome sequencing have given rise to models, such as the dynamic Flux Balance Analysis (dFBA) model, that allow us to apply more advanced control structures to bioprocesses due to their ability to account for a broad range of cellular behavior and operation conditions (Nakama & Jäschke, 2022).

The dFBA model consists of dynamic mass balance equations and an optimization that calculates the cell’s metabolic fluxes through the metabolic network. Thus, applying MPC results in a bi-level optimization problem, which is often very time-consuming to solve as we must analyze both the multiple upper-level optimization candidates as well as their corresponding lower-level candidates (Gupta et al., 2015).

The bi-level optimization problem can be avoided by reformulating the inner optimization with the KKT conditions, as shown by Ploch et al. (2020), which reformulated the dFBA into a differential-algebraic equation system. In order to embed the dFBA models into parameter optimization and optimal control problems, the system of equations can be discretized by applying orthogonal collocation with an adaptative mesh scheme (de Oliveira et al., 2023). The KKT reformulation of dFBA can also be used in model-based control, as shown by Nakama and Jäschke (2022).

However, these works have yet to explore the utilization of the duality theory reformulation, which is widely used in metabolic engineering (Zomorrodi & Maranas, 2013), for dFBA-based MPC. This work aims to apply MPC to a CSTR bioreactor based on the duality theory and KKT reformulations of the dFBA model of the *E. coli* core metabolic network and compare the solver time and reliability of the reformulations.

* 1. Methodology

The model developed in this work was based on the previous work of de Oliveira et al. (2023) and further extended using the duality theory approach and the CSTR mass balances. The purpose of this section is to provide a brief overview of the methodology used in that work. Readers are referred to the original publication for a more comprehensive description of the dFBA models.

* + 1. Case study

We consider a CSTR with *E. coli*, strain K-12, sub-strain MG1655, operating under aerobic growth with glucose and acetate. The feed flow consists of glucose and water, and the feed flow rate is calculated from the dilution rate of the reactor. We assume that oxygen is always available in the system, and we only consider the core metabolic network of the *E. coli* cell, which contains 95 metabolic reactions and 72 metabolites. The *E. coli* core metabolism is chosen because it is relatively simple and well-established. The metabolic data are gathered from the BiGG Models database (King et al., 2016).

* + 1. Flux Balance Analysis

The model is based on the Flux Balance Analysis (FBA) of the organism's metabolic network. The FBA is given as an optimization problem, and the objective is to maximize the cell's growth rate. To avoid multiple solutions, we use a variant of the FBA called the Parsimonious FBA (pFBA). The pFBA is expressed as presented by Ploch et al. (2020).

|  |  |
| --- | --- |
|  | (1) |

Here, **S** is the stoichiometric matrix for the metabolic reactions, **LB** and **UB** are vectors containing the lower and upper bounds for each reaction, **v** is a vector containing the metabolic fluxes, **c** is a vector containing weights that are multiplied with **v** to give the objective of the pFBA, and **W** is a diagonal matrix containing small weights, .

The pFBA is reformulated into algebraic expressions using the duality theory and KKT optimality conditions. The main difference between the two methods is how they implement the strong duality theorem. The duality theory approach directly implements the strong duality theorem by stating that the primal solution equals the dual solution. In contrast, the KKT approach replaces this expression with the complementary slackness conditions (Zomorrodi & Maranas, 2013). However, the reformulations may be difficult for the interior point solver to handle, partly due to the non-smooth characteristics of the system (Nakama & Jäschke, 2022). Therefore, we use a penalization method to relax the strong duality theorem constraints by reverting the reformulations back into optimization problems and moving the strong duality theorem constraints to the objective functions (Oliveira et al., 2023). The duality theory reformulation is given below.

|  |  |
| --- | --- |
|  | (2) |

Here *, ,* andare the Lagrangian multipliers*.* The KKT reformulation is given below.

|  |  |
| --- | --- |
|  | (3) |

* + 1. Dynamic FBA

The dFBA models assume that the intracellular reactions are much faster than the extracellular ones, therefore we consider the intracellular metabolites to be at quasi-steady state. The pFBA can thus be expanded to a dFBA by adding mass balances for the extracellular metabolites and kinetics for the uptake of the substrates. The dFBA is presented below.

|  |  |
| --- | --- |
|  | (4) |

Here D is the dilution rate, **x** is a vector containing the concentration of the extracellular metabolites, and are the maximal uptake and the Michaelis constant of the substrate i, respectively.

* + 1. MPC implementation

Before we apply the MPC, we first discretize the set of ordinary differential equations (ODE). It was decided to use orthogonal collocation as this technique is found to be equivalent to an implicit Runge-Kutta method, considering accuracy and stability (de Oliveira et al. 2023). We used three Radau collocation points, over a finite number of elements. Finally, we implement the MPC by expanding the objective function of the dFBA reformulations and by adding constraints for the manipulated variable, the dilution rate. The implementation of the MPC based on the penalized duality theory reformulation of the dFBA is presented below.

|  |  |
| --- | --- |
|  | (5) |

The MPC based on the KKT reformulation is presented below.

|  |  |
| --- | --- |
|  | (6) |

Here is the biomass concentration, is the biomass concentration setpoint, is the change in the dilution rate, is the maximal allowed change in the dilution rate, is the minimal value of the dilution rate, is the maximal value of the dilution rate, and Q, R, and C are tuning parameters for the MPC. The tuning parameters are found by trial and error, and the same tuning parameters are used for both the duality theory and the KKT MPC reformulations.

* + 1. Direct Approach for dFBA

We use the direct approach (DA) to develop a slightly different model of the system to use as a reference to our dFBA model reformulations and to act as the plant in our simulation with the MPC. In the DA model we apply an ODE solver that calls the pFBA model at each time step. The ODE solver calculates the metabolite concentrations and the lower bound of the substrates and provides them to our non-linear problem optimizer that solves the pFBA and returns the metabolic flux for each of the metabolites. We chose to use the DA as a reference because the DA solves the original pFBA problem and the adaptive step size utilized by the ODE solver makes the DA more accurate in areas with rapid change of the metabolite concentrations, given that the optimization problem returns feasible solutions at each time step. The ODE solver Quasi-constant time step Numerical Differentiation Function (QNDF) from the package DifferentialEquations (Rackauckas & Nie, 2017) is used, because the system is reasonably large, and the level of stiffness is unknown. The optimization problem is solved with the interior point optimizer IPOPT (Wächter & Biegler, 2016) with the MA97 linear solver from HSL (STFC, 2023).

* 1. Results and discussion

The MPC formulation based on the penalized duality theory (P. Dual) and the penalized KKT (P. KKT) are compared for a 35-hour simulation where the MPC was updated every hour and predicted 4 hours ahead of time. The simulation started at steady state and after 6 hours we increased the biomass setpoint with 50 % and reduced it back to the original value after 21 hours (Case 1). Figure 1 shows the biomass concentration and MPC solver time at each time step. We performed the same simulation for step changes in the glucose feed concentration (Case 2), and the maximal glucose uptake (Case 3), however we use a 50% decrease instead of increase in the glucose uptake parameter to avoid saturation.

A graph of a graph of a function

Description automatically generated with medium confidence

Figure 1: Biomass concentration (a) and computational time for each closed loop iteration (b) for simulation with 50 % increase in the biomass setpoint after 6 hours and a reduction back to the original value after 21 hours.

The average solver time, number of failed MPC optimizations, and the mean squared error (MSE) between the biomass concentration and the biomass concentration setpoint for the step changes in setpoint, maximal glucose uptake and glucose feed concentration are presented in Table 1.

Table 1: Average solver time, number of failed optimizations and mean squared error (MSE) for simulation with step changes in the biomass setpoint (Case 1), the glucose uptake (Case 2), and the glucose feed concentration (Case 3).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method: | Average solver time: [s] | Maximum number of iterations exceeded: | Failed: | MSE: |
| Case 1: | | | | |
| P. Dual | 0.62 | 0 of 35 | 0 of 35 |  |
| P. KKT | 4.93 | 11 of 35 | 2 of 35 |  |
| Case 2: | | | | |
| P. Dual | 0.74 | 0 of 35 | 0 of 35 |  |
| P. KKT | 4.64 | 10 of 35 | 0 of 35 |  |
| Case 3: | | | | |
| P. Dual | 0.52 | 0 of 35 | 0 of 35 |  |
| P. KKT | 2.91 | 8 of 35 | 0 of 35 |  |

From the simulationsit is apparent that the P. Dual reformulation is on averagemuch, roughly 5 to 8 times, faster than the P. KKT reformulation**.** However, when we look at Figure 1, we notice that the computational times are quite similar for most of the closed-loop iterations, except for 12 outliers. This aligns with the fact that 11 closed-loop iterations reached the maximum number of iterations, and the large average solver time for the P. KKT reformulation is likely due to these closed-loop iterations.

It is also observed that the optimization of the P. KKT reformulation is less likely to converge than the P. Dual reformulation. The two restoration failed messages for the setpoint change are particularly concerning since it indicates that the solver was unable to find a feasible point for these two optimizations. The higher reliability of the P. Dual may be attributed to its more linear characteristics, as its objective function contains fewer non-linear terms than the P. KKT reformulation. However, it is also possible that the control parameter weighting used in the two reformulations could be a contributing factor.

The MSE from the setpoint is relatively similar for the two approaches, which is likely due to most of the failed P. KKT optimizations being caused by the solver reaching the maximum number of iterations, thus failing to converge, but still getting very close to the true solution before the optimization is terminated.

* 1. Conclusion and future work

We have shown that MPC can be applied to CSTR bioreactors based on the P. Dual and the P. KKT reformulations of the dFBA. The MPC performed well and was able to keep the biomass concentration close to the desired setpoint and handle relatively large changes in the setpoint as well as disturbances in the glucose feed and in the maximal glucose uptake. Overall was the P. Dual reformulation found to be more reliable than the P. KKT reformulation when solved with the IPOPT solver. For future work the models need to be expanded to use a genome scale metabolic network of the *E. coli* and to be validated by experimental data.

* 1. Acknowledgements

This work was supported by the Research Council of Norway (RCN) through the FRIPRO Project SensPATH.

* 1. References

R. D. de Oliveira, G. A. C. Le Roux, and R. Mahadevan, 2023, Nonlinear programming reformulation of dynamic flux balance analysis models, Computers & Chemical Engineering 170, p. 108101.

P. M. Doran. 2013, Bioprocess Engineering Principles. 2nd ed, Academic Press.

A. Gupta, J. Mańdziuk, and Y. Ong, 2015, Evolutionary multitasking in bi-level optimization, Complex & Intelligent Systems 1, pp. 83–95.

B. Jabarivelisdeh et al., 2020, Adaptive predictive control of bioprocesses with constraint-based modeling and estimation, Computers Chemical Engineering 135, p. 106744.

Z. A. King et al., 2016, BiGG Models: A platform for integrating, standardizing and sharing genome-scale models, Nucleic acids research 44.D1, pp. D515–D522.

C. S. M. Nakama and J. Jäschke, 2022, Analysis of control models based on dFBA for fed-batch bioreactors solved by interior-point methods, IFAC-PapersOnLine 55.7, pp. 131–136.

T. Ploch et al., 2020, Simulation of differential-algebraic equation systems with optimization criteria embedded in Modelica, Computers Chemical Engineering140, p. 106920.

Science and Technology Facilities Council (STFC), 2023, A collection of Fortran codes for large scale scientific computation.

D. E. Seborg et al., 2016, Process Dynamics and Control. 4th ed., John Wiley & Sons.

C. Rackauckas and Q. Nie, 2017, Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia, Journal of Open Research Software 5.1.

A. Wächter and L. T. Biegler, 2016, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, Mathematical programming 106, pp. 25–57.

A. R. Zomorrodi and C. D. Maranas, 2013, Optimization methods in metabolic networks, John Wiley & Sons.