Optimal Feed Trajectories for Fedbatch Fermentation with Substrate Inhibition Kinetics

Thomas Emil Ryde, Morten Ryberg Wahlgreen, Marcus Krogh Nielsen, Steen Hørsholt, Sten Bay Jørgensen, John Bagterp Jørgensen

Technical University of Denmark & 2-control ApS DTU Compute & DTU Chemical Engineering DK-2800 Kgs Lyngby, Denmark.

Abstract: We provide optimal feed trajectories for fedbatch fermentation of microorganisms with substrate inhibition kinetics. We demonstrate that the optimal trajectories are non-unique and provide analytical procedures for solving the optimal control problem. Since, the optimal trajectories are non-unique this is essential for practical operations as a numerical solution could be any of the feasible non-unique optimal solutions. In addition, we investigate the sensitivity of the optimal trajectories and demonstrate that a bang-bang solution is optimal from a sensitivity point of view. This insight is relevant for model predictive control of fedbatch fermentation.

Keywords: Dynamic Optimization, Optimal Control, Fedbatch, Biotechnology, Process Control

1. INTRODUCTION

In biotechnology, fedbatch reactors are used for production of proteins (enzymes), beverages, and vaccines. Fedbatch operation is also used for start-up of novel continuously operated reactors such as continuous stirred tank reactors and the U-loop reactor. The results and insight provided by the optimal trajectories provided in this paper are relevant for optimal operation of such reactors and may also be relevant for semi-batch operation of chemical reactors.

Fedbatch bioreactors are widely used in the food, pharmaceutical and biotechnological industries due to the scale in such industries and their effectiveness in converting cells into biochemical products such as proteins (enzymes), beverages, and industrial solvents. Amongst the different operational techniques of a bioreactor, fedbatch operation has been particularly fruitful due to its economic relevance (Lim and Shin, 2011). Fedbatch operation is an operational technique, where substrate is fed to the bioreactor without removing product until the bioreactor is full.

In fedbatch operation of a bioreactor, the aim typically is to maximize the biomass yield or the yield of a certain product, which typically is strongly related to the biomass growth. Therefore optimal control of a bioreactor is of great importance for process engineers. However, optimal control of a bioreactor has proven to be challenging due to its very nonlinear behaviour and absence of steady states. Despite advances in computational power and numerical algorithms for optimal control problems implemented in a receding horizon manner, surprisingly, the use of such optimization based feedback enabled techniques are still not widely used for control of fedbatch operated bioreactors (Engell, 2007). Furthermore, important process variables can not always be measured which contributes to making making feedback control difficult in practice (Vojinović et al., 2006). To understand if there is a deeper fundamental reason for the lack of application of numerical nonlinear model predictive control for fedbatch operation of fermentors, we derive analytic solutions to nonlinear optimal control problems. These analytical solutions provide structural insight of the optimal trajectories and in particular demonstrate that the optimal control problem has non-unique optimal solution trajectories. This insight is important for the numerical solution of such optimal control problems. Regularization and homotopy techniques based on this insight can be used to steer the numerical solution towards a particular solution that is desirable from a practical point of view.

Optimal operation of a bioreactor using a simple model is often subject to manipulating a substrate feed rate and a water inlet variable. Determination of the feed rate often results in a singular control problem, due to the control variables appearing linearly in the Hamiltonian, such that Pontryagin's maximum principle fails to provide a solution. Weigand (1981) used a mass balance model to solve a repeated fedbatch fermentation process. To overcome the issue of singular control, they use the second order derivative to produce conditions which yield optimal control. More recently, Petersen and Jørgensen (2014) used a similar mass balance model with two liquid inlets instead of a substrate feed and a water inlet. This variable choice (variable transformation) allowed them to derive an analytic solution. Using the same model as Petersen and Jørgensen (2014), we show that optimal trajectories in the mass balance model are non-unique. We provide methods to find optimal trajectories and use in silico simulations to investigate the sensitivity of the solutions.



Fig. 1. A sketch of a bioreactor in fedbatch operation. The diagram shows the model state variables and control inputs.

This paper is organized as follows. We present the fermentation model and the optimal control problem in Section 2. Section 3 and 4 derive analytical expressions for 3 selected optimal control trajectories, and provide a numerical sensitivity study of these 3 selected optimal control trajectories. Section 5 concludes that the optimal control problem for fedbatch operation of microorganism with Haldane growth has a non-unique optimal solution. Sensitivity analyses reveal that the bang-bang solution is least sensitive to the initial conditions and stochastic process noise.

2. FERMENTATION MODEL

We derive a mathematical model that describes a fedbatch fermentation process using stoichiometric and kinetic parameters for *Methylococcus Capsulatus* (Olsen et al., 2010). Table 1 provides the stoichiometric, kinetic and operational parameters for this fedbatch fermentation. As illustrated in Fig. 1, the model variables are the volume, V, the biomass concentration, c_X , and the substrate concentration, c_S , along with water inlet rate, F_W , and substrate inlet rate, F_S . We use Haldane growth kinetics to model the reaction from substrate to biomass and assume that the system is oxygen saturated.

We consider a differential equation of the following form,

$$\frac{d}{dt}g(x(t)) = f(x(t), u(t)), \qquad t \in [t_0, t_f],$$
(1)

where $x(t_0) = x_0$. This form is natural in many applications, namely within process engineering. g(x(t)) typically represents mass, momentum, or energy. Assuming that dg/dt is non-singular, one often rewrites the above form to,

$$\frac{d}{dt}x(t) = \left(\frac{dg}{dx}x(t)\right)^{-1}f(x(t), u(t)), \qquad t \in [t_0, t_f], \quad (2)$$

where $x(t_0) = x_0$. However, this form is not numerically equivalent to the form in (1) and does not guarantee conservation of mass, momentum, energy, etc. Thus, we keep the model in the form given by (1).

Table 1. Parameters for the fedbatch fermen-
tation of M. Capsulatus.

Variable	Value	Unit
μ_{max}	0.37	1/h
K_S	0.021	kg/m^3
K_I	0.38	kg/m^3
γ_s	1.777	kg substrate/kg biomass
$c_{S,in}$	10.0	kg/m^3

2.1 Stoichiometry and kinetics

The reaction stoichiometry is given by,

$$\gamma S \to X,$$
 (3)

where γ denotes the stoichiometric coefficient. The reaction rate governed by Haldane growth kinetics is given by, $r = \mu(a_{\tau})a_{\tau}$ (4)

The specific growth rate,
$$\mu(c_S)c_X$$
. (4)

$$\mu(c_S) = \mu_{max} \frac{c_S}{K_S + c_S + c_S^2 / K_I},$$
(5)

where c_X and c_S respectively denote the biomass and substrate concentration. Thus, the reaction rate is limited by substrate inhibition. The production rate of biomass, X, and substrate, S, is defined from the stoichiometry in (3),

$$R_X = r, \qquad R_S = -\gamma r. \tag{6}$$

2.2 Mass balances

The manipulated variables are the inlet flow rates, i.e. the substrate inlet flow rate, F_S , with a fixed substrate concentration, $c_{S,in}$, and the water inlet flow rate, F_W . The total flow rate, F, to the reactor is

$$F = F_W + F_S \tag{7}$$

and the corresponding substrate concentration is \overline{D}

$$c_{in} = \frac{F_S c_{S,in}}{F} = \frac{F_S c_{S,in}}{F_W + F_S}.$$
 (8)

The fermentation process in fedbatch operation is modeled using mass balances, yielding the following system of differential equations,

$$\frac{d}{dt}(V) = F_S + F_W,\tag{9a}$$

$$\frac{d}{dt}(Vc_X) = R_X V, \tag{9b}$$

$$\frac{d}{dt}(Vc_S) = F_S c_{S,in} + R_S V, \tag{9c}$$

where $V(t_0) = V_0$, $c_X(t_0) = c_{X,0}$, and $c_S(t_0) = c_{S,0}$. Each of the three equations respectively model the total volume, V, the total mass of biomass, $m_X = Vc_X$, and the total mass of substrate, $m_S = Vc_S$.

2.3 Objective function

The objective of the fedbatch fermentation is to maximize the produced biomass. This can be expressed by the objective function,

$$\phi = \int_{t_0}^{t_f} R_X V dt, \qquad (10)$$

which represents the produced biomass within the time interval $[t_0, t_f]$.

 Table 2. Operational restrictions on control inputs and state variables.

Variable	Value	Unit
Vmax	12.39	m^3
$c_{X,max}$	2.00	kg biomass $/m^3$
$c_{S,max}$	3.00	kg substrate/ m^3
$F_{s,max}$	10.00	m^3/h
$F_{w,max}$	10.00	m^3/h

2.4 Input constraints and path constraints

The fermentation process is subject to operational constraints. The input constraints are given by,

$$0 \le F_S \le F_{S,\max}, \tag{11a}$$

$$0 \le F_W \le F_{W,\max}. \tag{11b}$$

$$0 \leq F_W \leq F_{W,\max}$$

The state constraints are given by,

$$0 \le V \le V_{\max},\tag{12a}$$

$$0 \le c_X \le c_{X,\max},\tag{12b}$$

$$0 \le c_S \le c_{S,\max},\tag{12c}$$

where $c_{X,max}$ and $c_{S,max}$ are chosen to keep the states in a region that ensures model accuracy. Table 2 shows the parameters for the constraints.

3. DERIVATION OF OPTIMAL TRAJECTORIES

In this section we derive selected optimal control trajectories for the optimal control problem presented in Section 2. We derive optimal trajectories that maximize the produced biomass over the fixed time period $[t_0, t_f]$.

Let $\beta(t) = R_X V$ denote the integrand in the objective function (10) and $\beta^*(t)$ the optimal trajectory of $\beta(t)$. Petersen and Jørgensen (2014) showed that,

$$\beta^*(t) = \mu(c_S^*) c_{X,max} V_0 \exp(\mu(c_S^*) t).$$
(13)

As the specific growth rate, $\mu(c_S)$, has a maximum, $\mu(c_S^*)$, we consider solutions where $c_S^* = \arg \max \mu(c_S) = \sqrt{K_I K_S}$. If $\beta(t) = \beta^*(t)$, we require that,

$$c_X V = c_{X,\max} V_0 \exp(\mu(c_S^*)t).$$
 (14)

At $t = t_0$ (14) implies that the initial biomass is given by, $c_{X,0} = c_{X,\text{max}}.$ (15)

Inserting (9a) into (9c) and utilizing (14) we get,

(

$$\frac{a}{dt}(V)c_S^* = F_S c_{S,in} + R_S V \tag{16a}$$

$$(F_S + F_W)c_S^* = F_S c_{S,in} - \gamma R_X V \tag{16b}$$

$$F_S + F_W)c_S^* = F_S c_{S,in} - \gamma \beta^* \tag{16c}$$

$$F_S = \frac{F_W c_S^* + \gamma \beta^*}{c_{S,in} - c_S^*}.$$
 (16d)

To ensure feasibility, we must show that this control input does not violate state constraints. The lower bounds on the states are automatically satisfied due to the non-negativity of the inputs. We consider the upper bound of (12a). We need to show that,

$$V \le V_{\max}.\tag{17}$$

As V is non-decreasing, it is sufficient to consider whether the total volume V at $t = t_f$ exceeds V_{max} . We integrate both sides of (9a) to get,

$$V_{\max} = V_0 + \int_{t_0}^{t_f} \left(F_S + F_W \right) dt.$$
 (18)

We integrate (9c) and apply (13) to obtain the total inlet from F_S ,

$$c_{S}^{*}(V_{\max} - V_{0}) = \int_{t_{0}}^{t_{f}} (F_{S}c_{S,in} + R_{S}V) dt$$
(19a)

$$\frac{c_S^*(V_{\max} - V_0)}{c_{S,in}} = \int_{t_0}^{t_f} \left(F_S - \frac{\gamma}{c_{S,in}} \beta^*(t) \right) dt \qquad (19b)$$

$$\int_{t_0}^{t_f} F_S dt = \frac{(c_S^* + \gamma c_{X,\max})(V_{\max} - V_0)}{c_{S,in}}.$$
 (19c)

Consequently, the total inlet from F_S is constant. Similarly, using (18) and (19c), the total inlet from F_W is constant,

$$\int_{t_0}^{t_f} F_W \, dt = \frac{(c_{S,in} - c_S^* - \gamma c_{X,\max})(V_{\max} - V_0)}{c_{S,in}}.$$
 (20)

Thus if (20) is satisfied, F_S , given in (16d), satisfies the upper bound (12a). As the total input from F_W is non-negative, we note that (20) implies,

$$c_{S,in} \ge c_S^* + \gamma c_{X,\max}.$$
 (21)

Now we consider the upper bound of (12b). We need to show that,

$$c_X \le c_{X,\max}.\tag{22}$$

Using (14) and the integral of (9a) from $t' = t_0$ to t' = t, we need to show that,

$$c_X \le c_{X,\max} \tag{23a}$$

$$\frac{c_{X,\max}V_0\exp(\mu(c_S^*)t)}{V} \le c_{X,\max}$$
(23b)

$$V_0 \exp(\mu(c_S^*)t) - V_0 \le \int_{t_0}^t F_S(t') + F_W(t') \, dt'. \quad (23c)$$

From (16d), we find that,

$$\int_{t_0}^t F_S(t') dt' = \int_{t_0}^t \frac{F_W(t')c_S^* + \gamma\beta^*(t')}{c_{S,in} - c_S^*} dt'.$$
 (24)

The right hand side of (24) is bounded by applying (21),

$$\frac{\gamma}{c_{S,in} - c_S^*} \int_{t_0}^t \beta^*(t') \, dt' = \frac{\gamma(\beta^*(t) - \beta^*(t_0))}{\mu(c_S^*)(c_{S,in} - c_S^*)} \tag{25a}$$

$$\leq \frac{\gamma(\beta^*(t) - \beta^*(t_0))}{\mu(c_S^*)(\gamma c_{X,\max})} \tag{25b}$$

$$= V_0(\exp(\mu(c_S^*)t) - 1). \quad (25c)$$

We now insert the right hand side of (24) into the right hand side of (23c). Utilizing (25c), we find that,

$$0 \le \frac{c_S^*}{c_{S,in} - c_S^*} \int_{t_0}^t F_W(t') \, dt' + \int_{t_0}^t F_W(t') \, dt' \qquad (26a)$$

$$\leq \left(1 + \frac{c_S^*}{c_{S,in} - c_S^*}\right) \int_{t_0}^t F_W(t') \, dt'.$$
 (26b)

As the right hand side of the above is positive, we conclude that (22) is satisfied and thus the upper bound of (12b) is not violated. Lastly, we consider the upper bound of (12c). This bound is satisfied by choosing $c_{S,\max}$ such that $c_S^* \leq c_{S,\max}$.

We have shown that the optimal control problem is solved with any control input that satisfies (16d) and (20), and does not violate the operational restrictions given by (11a) and (11b). Thus, the solution to the optimal control problem is non-unique.



Fig. 2. The control input for three different solutions to the deterministic optimal control problem.

4. OPTIMAL CONTROL AND SENSITIVITIES

In this section, we present three concrete solutions to the optimal control problem that satisfy the previously derived conditions. We also investigate the sensitivity of the solutions using Monte-Carlo simulation and stochastic differential equation extensions of the model.

The first solution fixes $R_X = R_{X,\max} = \mu(c_S^*)c_{X,\max}$ throughout the operation. As a result, we have $R_S^* = -\gamma_s R_{X,\max}$. We name this solution *Maximum Biomass Concentration*. In the second solution we fix $F_W = F_{W,\max}$ until a switching time, $t = t_{switch}$, and then $F_W = 0$. Using (20) one can compute t_{switch} ,

$$t_{switch} = \frac{(c_{S,in} - c_S^* - \gamma_s c_{X,max})(V_{max} - V_0)}{F_{w,max} c_{S,in}}.$$
 (27)

We name this solution *Bang-Bang*. In the third solution F_W is computed from (20) and kept constant throughout the operation. We name this solution *Constant Water Flow Rate*. Fig. 2 provides the control inputs for the three control strategies that are all optimal, i.e. the deterministic optimal control problem has the same value for all 3 input profiles. Fig. 3 illustrates the corresponding trajectories.

Fig. 4 shows the state variables for the three solutions. In all three solutions c_S diverges from c_S^* , which indicates that a sensitivity study is of interest. We use Monte-Carlo simulations and stochastic simulations to compare the sensitivities of the 3 different optimal solutions.

4.1 Perturbations in initial conditions

We first consider normally distributed initial conditions. We consider the case where $c_{S,0} \sim \mathcal{N}(c_S^*, 0.02^2)$ and $c_{X,0} \sim \mathcal{N}(c_{X,\max}, 0.1^2)$. We run 5000 simulations for each

Table 3. Nominal initial conditions.

Value	Unit
1.00	m^3
2.00	kg biomass $/m^3$
0.0893	kg substrate/ m^3
	Value 1.00 2.00 0.0893



Fig. 3. The produced biomass and inputs to the bioreactor when using the three chosen control strategies. The three solutions produce the same optimal amount of biomass.

perturbation and compute the produced biomass at $t = t_f$. Note that a simulation is not stopped if $c_X \ge c_{X,\max}$, i.e this enables the produced biomass at $t = t_f$ to exceed the case with deterministic initial conditions. Fig. 5 shows the trajectory of 10 of these 5000 simulation for $c_{X,0} \sim \mathcal{N}(c_{X,\max}, 0.1^2)$.

Fig. 6 shows that the Bang-Bang control strategy leads to the largest 10th quantile and mean produced biomass at time $t = t_f$ for both perturbations in initial conditions. Based on these simulations, we conclude that the *Bang-Bang* control strategy is least sensitive (more robust) to the given perturbations in initial conditions. This makes intuitive sense as the *Bang-Bang* solution fills the bioreactor as fast as possible, making a small difference in $n_{X,0}$ or $n_{S,0}$ have less impact on c_X and c_S shortly after operation start.

4.2 Process noise

We now consider the impact on produced biomass when adding a stochastic diffusion term to the deterministic differential equation model (9). The resulting model is represented as a system of stochastic differential equations (SDEs),



Fig. 4. State variables of the bioreactor when using the three chosen control strategies. The *Bang-Bang* solution and the *Maximum Biomass Concentration* solution respectively define the upper and lower bound of the speed which the bioreactor is filled. c_S diverges from c_S^* in all three solutions. This indicates that the optimal trajectory is a repeller.

$$dV = (F_s + F_w) dt + \sigma_1 dB_1, \qquad (28a)$$

$$d(c_X V) = (R_X V) dt + \sigma_2 dB_2, \qquad (28b)$$

$$d(c_{S}V) = (F_{s}c_{S,in} + R_{S}V) dt + \sigma_{3}dB_{3}, \qquad (28c)$$

where $V(t_0) = V_0$, $c_X(t_0) = c_{X,0}$, and $c_S(t_0) = c_{S,0}$. Furthermore, B_k denotes a standard Weiner process, i.e. $dB_k(t) \sim N(0, dt)$. We consider two types of noise, case 1 with a constant diffusion coefficient across all three equations, and case 2 where the diffusion coefficient scales with the size of the drift term,

$$\begin{array}{ll} (1) \ \, \sigma_1=\sigma_2=\sigma_3=C, \\ (2) \ \, \sigma_k=\sqrt{C|f_k|}, & k\in\{1,2,3\}, \end{array}$$

where $C \in [0.02, 0.2]$ is a constant and f_k denotes the drift term for equation k. We determine the distribution of the produced biomass at $t = t_f$ for the three control strategies for different values of C. We use the Euler-Maruyama numerical scheme with $\Delta t = 10^{-4}$ to solve 500 realization of the SDE for each control strategy and each value of C.

Fig. 7 and Fig. 8 illustrate the produced biomass as function of the diffusion scaling coefficient. It is evident that the *Bang-Bang* solution results in the largest mean and 10th quantile of the produced biomass for all values of C and both types of noise. In Fig. 7 we observe that even the 10th quantile of the produced biomass when using the *Bang-Bang* solution exceeds the mean produced biomass when using the *Maximum Biomass Concentration* solution for all values of C. Thus, it seems that the *Bang-Bang* solution is the most robust to both presented noise types.



Fig. 5. The trajectory from 10 draws from the distribution $c_{X,0} \sim \mathcal{N}(c_{X,\max}, 0.1^2)$. Amongst these 10 draws, we see that the *Bang-Bang* solution results in most produced biomass. The dotted black trajectory represents the solution without the presence of noise. All other colored trajectories correspond to the 10 draws.

5. CONCLUSION

In this paper, we consider substrate and water feeding profiles that maximize productivity of a bioreactor in fedbatch operation. We model the bio-culture dynamics using a mass balance model and a reaction rate based on Haldane growth kinetics. We show that the solution of the optimal control problem is non-unique, and that the optimal control problem is solved with any control input that satisfies (16d) and (20), and does not violate the operational restrictions given by (11a) and (11b).

In addition to deriving optimality conditions, we investigate the sensitivity of three concrete solutions to the optimal control problem. We name these solutions the *Maximum Biomass Concentration* solution, the *Bang-Bang* solution and the *Constant Water Flow Rate* solution. We investigate solution sensitivity by Monte-Carlo simulations of the initial conditions and simulations of equivalent stochastic differential equations with process noise. Specifically, we used open-loop simulations to investigate the impact on produced biomass when perturbing the initial conditions and when adding model randomness using a system of stochastic differential equations (SDE). Our numerical simulations show that the *Bang-Bang* solution is least sensitive to the perturbations, and results in a larger



Fig. 6. The distribution of n_X at $t = t_f$ in the presence of perturbations in the initial conditions. Left: $c_{X,0} \sim \mathcal{N}(c_{X,\max}, 0.1^2)$. Right: $c_{S,0} \sim \mathcal{N}(c_S^*, 0.02^2)$.



Fig. 7. The produced biomass, n_X , as a function of the diffusion coefficient, $\sigma = C$, at $t = t_f$ for the three control strategies. The non-dotted line represents the mean produced biomass, while the dotted line represents to the 10th quantile for the corresponding control strategy.

biomass yield in the presence of the investigate types of randomness.

The insight provided by the analytical optimal feed profiles is very important for nonlinear model predictive control of



Fig. 8. The produced biomass, n_X , as a function of the diffusion coefficient, $\sigma_k = \sqrt{C|f_k|}$, at $t = t_f$ for the three control strategies. The non-dotted line represents the mean produced biomass, while the dotted line represents to the 10th quantile for the corresponding control strategy.

such fedbatch fermentations; and may explain the limited use of nonlinear model predictive control in *industrial* fedbatch fermentations. Numerical solution of the optimal control problem in nonlinear model predictive control of such fedbatch fermentations is not well defined in the sense that the optimal solution is non-unique and the optimization algorithm may converge to any of the optimal solutions. It is not guaranteed that it will converge to the bang-bang solution that is least sensitive to additive process noise and initial condition perturbations. While the sensitivity can be reduced by feedback enabled control of the substrate, the bang-bang nominal trajectory is still expected to be the most robust.

Further work includes implementation of nonlinear model predictive control for simulated case studies of fedbatch fermentation, and implementation in an industrial fedbatch fermentor.

REFERENCES

- Engell, S. (2007). Feedback control for optimal process operation. *Journal of Process Control*, 17(3), 203–219.
- Lim, H. and Shin, H.S. (2011). Fed-batch cultures: Principles and applications of semi-batch bioreactors. Cambridge University Press.
- Olsen, D.F., Jørgensen, J.B., Villadsen, J., and Jørgensen, S.B. (2010). Modeling and simulation of single cell protein production. *Computer Applications in Biotech*nology, 11(1), 502–507.
- Petersen, L.N. and Jørgensen, J.B. (2014). Real-time economic optimization for a fermentation process using model predictive control. *European Control Conference*, (1831-1836).
- Vojinović, V., Cabral, J.M.S., and Fonseca, L.P. (2006). Real-time bioprocess monitoring: Part i: In situ sensors. Sensors and Actuators, B: Chemical, 114(2), 1083–1091.
- Weigand, W.A. (1981). Maximum cell productivity by repeated fed-batch culture for constant yield case. *Biotechnology and Bioengineering*, 23(2), 249–266.