Optimization of Metabolic Networks for Metabolite Overproduction

Creutzburg, K.; Jørgensen, Sten Bay

Publication date:
2001

Document Version
Publisher's PDF, also known as Version of record

Optimization of metabolic networks for metabolite overproduction

Kurt Creutzburg

Computer Aided Process Engineering Center
Department of Chemical Engineering
Technical University of Denmark

November 29, 2000

ISBN 87-90142-77-2
Printed by Bookpartner, Nørhaven Digital, Copenhagen, Denmark

Preface

This work is done as a partially fulfillment of the requirement for the PhD-degree at the Technical University of Denmark. The work was carried out at the Department of Chemical Engineering with Professor Sten Bay Jørgensen as supervisor. The project was performed over the period September 1, 1995 to April 30, 1999 and financed by the Danish Technical Research Council.

Acknowledgments

I would like to thank my family and friends for their encouragement and patience during this work. I would like to thank my supervisor Professor Sten Bay Jørgensen for making it possible for me to do my Ph.D. project at department of Chemical Engineering and for interesting discussions during these last years. This work have been sponsored by the Danish Technical Research Council, which is gratefully acknowledged.

I would like to thank Professor Lengeler for providing me insight on the thrilling world of biochemistry and gene regulation, during my stay in Stuttgart in the fall of 1996.

It has always been a pleasure to work with the members of the CAPEC with whom I had many interesting discussions, on the biotechnical issues with Teresa Zangirolami, Lars Gregersen, Frede Lei, Anders Nordstrøm and on nonlinear analysis fruit full discussions with Bodil Recke and Britta R. Andersen. A number of people have provided inputs for control ideas and generated good humor, Torben R. Andersen, John Bagterp, Kalvs Esbjerg, Hassan Yazdi, Jens Erik Hansen, Jens Meinhold, Ulrich Krühne and Alain Larose.

Finally a special thanks goes to my dearest Kamille for her patience and understanding, also thanks to Andreas and Mads for generating good spirit in the few hours after work.

Kurt Creutzburg
Year 2000

Resume

Driftoperationen af græinger kan gennemføres under meget forskellige vilkår. Dette betyder at forholdene for maximal udnyttelse af resourcer er meget forskellige, disse forhold er undersøgt for at sikre optimale drift betingelser.

Der gennemføres analyse af den kontinuert gærings proces' opførsel ved kritisk fortyndingshastighed. Dette gøres med to forskellige metoder. Der gennemføers model identifikation vha data opsamlet omkring kritisk fortyndingshastighed. Dette gøres for at undersøg om en to trins identifikations metode designet til analyse af lukket sløjfe proces data vil forbedre model identifikationen. Da maximal biomasse produktion af gæringen sker i et område hvor der er flere stationære tilstande genemføres operationen i lukket sløjfe. Analysen afslører at procesen er ikke lineær da tidskonstanterne af de identificerede modeler er afhnig af operations punktet. Der er dog ikke noget endegyltigt bevis for ustabile stationære tilstande, da de identificerede proces modeler ikke har positive egenværdier.

Den anden angrebsvinkel til problemstillingen er at analysere en matematisk proces model af gæringsprocesen vha ikke lineær model analyse for at afdække operationelle problemer. Ligeledes undersøges samspillet mellem fortyndingshastighed og substrat koncentration set ud fra en stabilitetsbetragtning.

Analysen af proces modelen viser at der er multiple stationære tilstande omkring den kritiske fortyndingshastighed og at der er en stærk korrealtion mellem fortyndingshastighed og substrat koncentration.

Baseret på den ikke lineære analyse foreslås en ny regulerings metode, der testes via simulering. Det viser sig meget lovende.

Denne afhandling kombinere områderne indenfor bioteknologi, ikke lineær model analyse og proces regulering. Ved analyse af den bioteknologiske proces model findes kvalitativt en opførsel svarende til eksperimentelle data. Typisk for gærings procesen er at der ikke er mange muligheder for regulering og kun få målinger som ikke nødvendigvis er direkte anvendelig til regulering. Der er også proces variable der ikke er let tilgængelig, som biomasse måling.

Et interessant punkt ved udviklingen af regulerings teorien igennem de sidste halve århundrede er at enten er det baseret på linear teori og derfor ikke specielt velegnet til at løse ikke lineære problemer eller baseret p ikke linær teori hvor analysen og regulator design er kompliceret af meget tung matematik, det bevirker at det kun er anvendelig på mindre systemer. Verden er jo klart ikke lineær, den biokemiske verden er ydermere kompliceret af det store antal kemiske komponenter der er involveret.

Summary

The operation of yeast cultivations is the main focus of this work and since the operation can be conducted in significantly different modes these are investigated to assure operation near optimal conditions.

The analysis of fermentation process behaviour is performed by two approaches. First a set of experimental data obtained in the region around the critical dilution rate of a continuous cultivation will be analysed to investigate if something is to be gained by applying novel identification techniques. Since the operation is conducted in the region where multiple steady states are expected, the experiments are performed in closed loop. The result of this analysis is that clearly the process is nonlinear since the time constants of the process changes (not superposition principle for linear systems). But the proof of the unstable steady states (negative eigenvalue) was not found in these set of experiments.

The second approach is to investigate a mathematical model of yeast fermentation by applying nonlinear process analysis techniques for determining the possibilities for operational problems and determine the stability of different operation points. The analysis focuses on the influence of the variables which is available for control (Dilution rate and substrate concentration)

The analysis of the proposed model show that there exists multiple steady states in the region around the critical dilution rate and the region is strongly dependent upon the substrate concentration.

Based on the nonlinear analysis a novel control scheme is suggested and tested in a number of simulation studies.

The strategy shows to be promising

This work have combed the fields of biotechnology, nonlinear process analysis and process control. By examining how a biological process behaves, the aim is to operate the process such that the desired behaviour is achieved. The typical biotechnological process does not have many control handles and only and a few measurements, which is not necessarily what is needed for the control, by this i mean, as an example, the measurement of the biomass concentration can not be measured directly on-line. There are developed techniques that can make the measurement of some intermediate possible, e.g. quenching techniques, though not developed for application in large scale industrial fermentation vessels, but very applicable in laboratory facilities. Another issue is that the control theory that have been developed over the past half century is either linear and thereby limited from solving nonlinear control problems but simple and attractive from a mathematical point of view or nonlinear control theory which involves much more complicated mathematics, which is hard to solve for even medium size problems. Clearly the world is nonlinear and especially biochemical processes are by the interaction of the many substrates,

enzymes, genes and products.

To improve the understanding of the underlying transitions causing the appearance of multiple steady states nonlinear model process analysis is performed.

Contents

1

Introduction

1.1 Background

The motivation for this work is the reported evidence of multiple steady states in continuous fermentation processes. The existence of multiple steady states in fermentation processes is a result of nonlinearities in the process. Non linearities are known to cause problems in the chemical process industry ranging from operational difficulties to serious safety problems.

The safety problems reported for chemical process due to nonlinear behaviour, e.g. runaway of temperature and pressure in reactions, is not considered as a problem here. The effect of multiple steady states in continuous fermentation processes, will result in loss of productivity.

Since the fermentation process is generally regarded as a safe process due to the low temperatures and moderate pressures, the problems generated by multiple steady states is operational. Actually the major safety problem is cleaning the vessel between fermentations and sterilising media, due to high temperature.

The multiple steady states are only present in a region around the critical dilution rate. The phenomena observed is two fold bifurcations resulting in three steady states of which one is unstable. In order to ensure stable operation near optimal operational conditions (close to critical dilution rate), feedback control is introduced. The closed loop operation can be used for experimental investigation of the operation around the critical dilution rate, both in stable and unstable regions.

The basic fermentation process has been known for thousands of years, areas as bread making, brewing and wine making, but the last half century have developed whole industries which apply fermentation for a wide range of products, such as pharmaceutical products, industrial enzymes, food ingredients, detergent etc. The development of these industries have placed focus on understanding how the microorganisms behave given certain conditions.

One question to ask is why are the multiple steady states not reported as a major problem in industrial fermentation processes ?

There are several answers to this a few can be

- Industrial secrecy
- The industrial fermentation process are run at conditions where multiple steady states does not occur.

The biotechnological research has reached a level of understanding where it is relevant to model the metabolic pathways within the given microorganism. The modelling of a given microorganism must reflect it's purpose. If the model is to be used in a study for increasing the knowledge of the microorganism functionality the level of detail has to be sufficient to reflect this purpose. Some flux analysis models have been developed with as many as 100 components. On the other hand if the model is to be used for control and operational investigations the necessary and sufficient level of detail might not be as high, e.g. substrate and biomass and maybe a by product. The trade off is between the complexity/information and simplisity/overview has to be made at the initial stage when selecting modelling strategy.

One could ask why the interest in yeast, *Saccharomyces cerevisiae*?. The simple answer is that many functions within the cell is not fully understood. The network of interactions between substrates, genes, proteins is very complex and not understood in detail, this include regulatory aspects on different levels. Historically the organism has been used for bread making and brewing for centuries, so the experience with the microorganism is well documented. In the last half century the development in the pharmaceutical industry have been to use microorganisms as hosts for the production of complex compounds by insertion of appropriate genes. The next question could be why then model such a well tried out organism? The question can be answered with this, the organism have been applied for many purposes but many details are still not understood even though the gene have been sequenced, to help the understanding of the relationships in the cell the modelling can be use as a tool. Models of fermentation processes are also useful in the analysis of process behaviour and control design and analysis.

1.2 Thesis Structure

The goal of this thesis is to:

- Found the basis for linking genetic/biochemical knowledge to functional behaviour.
- Investigate fermentation processes from an operational point of view to improve understanding of the process around the critical dilution rate and increase productivity.
- Apply new tools for process data analysis and process model analysis to reveal new insight of the fermentation process.

To address these goals the thesis is divided as follows: Chapter 2 describes different modelling approaches depending on the use, which reveals in the level of detail of these model types. Chapter 3 is an introduction to a mathematical model of yeast from and operational point of view. The model describes the relations between the components in yeast fermentation, biomass, substrate and

ethanol. The description of the model will be used in subsequent chapters. Chapter 4 is a study of a series of continuous fermentation experiments perform in closed loop (with feedback control). Here a novel technique is applied for analysis of closed loop experimental data. Chapter 5 and 6 is model analysis of the model described in chapter 3 and statement of the control problem for processes with multiple steady states, including a solution. Chapter 7 is an investigation of optimization of on a penicillin cultivation with dynamic optimization of feeding profile. In chapter 8 and 9 the overall conclusion and a discussion on further directions in this field is pointed out.

There are three appendices following this thesis Appendix A is a brief introduction on bifurcation theory with two minor examples for introducing the terminology. There are also an introduction to the algorithm which is used for the model analysis. Appendix B present a linearization of the model presented in chapter 3 this linearized model is use for bifurcation purposes and calculation eigenvalues. Appendix C is a paper written on controlability the techniques presented in the paper are similar to the tools which are used in optimization techniques to assure the process is controlled within a specified range. The tools applied for controlability analysis of nonlinear process models nonlinear optimization problems similar to the method applied in Chapter 7.

2

Modelling concepts for microorganisms. An introduction

This chapter will introduce the background for the later chapters on biochemical issues. The introduction is mainly given for readers not familiar with the field of biotechnology. To be able to explain the nonlinear phenomena in the biochemical processes a necessary prerequest is an understanding on the mechanisms of the metabolism in the microorganism.

2.1 Introduction

In the later years many researchers have been studying special parts of the metabolism of different microorganisms. The most interesting organisms from an industrial point of view are *Eschericia coli* and *Saccharomyces cerevisiae*, since they are widely used and reasonable well understood. These two organisms are used for many purposes either as host organisms for protein production or as the biomass product. Researchers are working on different levels of the metabolic structure of the organisms, some on the genetic level, some at the biochemical level and some at the reactor level. One limitation to gain full advantage of the information is the lack of a common language between the different researchers in this area. One way to structure the available information on the organisms is through mathematical modelling of the whole system, by this approach there would be generated a library of information ranging from genetic information to reactor performance. The work done in the EcoCyc project Karp et al. (1999) provides some of the information for *E. coli*.

When looking at the whole system, the details of some parts might be negligible whereas some other parts play a dominant role. One formulation which is appropriate is the citation in Bailey (1998), which is taken from Casti (1992a) and Casti (1992b), it goes as follows; *Basically, the point of making models is to be able to bring a measure of order to our experience and observations, as well as to make specific predictions about certain aspects of the world we experience*. Another reference in this context is Nørretranders (1991) where

there is a discussion on the details on a map (in this context it is road maps) how large should a map be to describe all details?, the answer is that the map is equal to the whole. The map needed for driving on highways needs not so many details as for driving in the countryside. The same guidelines are valid for mathematical modelling of biochemical processes, the level of detail should reflect the area of application. Again from Bailey (1998) *Mathematical modelling does not make sense without defining, before making the model, what its use and what problem its intended to help to solve.*

One could say that the field of modelling the biochemical processes constitutes a link between the biochemists/micorobiologist and the biochemical/bioprocess engineer, without an interplay between these two groups it is hard to produce reliable models for any purpose.

The intention with this chapter is to give an overview of different modelling types and an insight to the application of the different types of models

In this chapter there is a description of some observed features of microorganisms and some of the preliminary tools to analyse the metabolism of the cell. To understand the behaviour of the cell and the reasoning in subsequent chapters it is necessary to understand how cells behave and react to changes in the surroundings.

The chapter is structured as follows, first there is a brief description of some fundamental behaviour observed for microorganisms in different types of operation in section 2.2.

To understand the mechanisms that control the fluxes in a microorganism it is necessary to understand the basic biochemistry which holds the overall control of the enzyme machinery needed for the metabolic reaction network. Concepts such as operons and regulons are explained in section 2.3, these biological control systems are responsible for the control at the genetical level.

A discussion on the modelling approach in unstructured/structured biomass models is given in section 2.4 and 2.5.

Then in section 2.5.1 flux analysis and metabolic control analysis are introduced with a short description of the application of such modelling and analysis techniques for engineering purposes.

2.2 Experimental knowledge

In the literature on of the most quoted mathematical model of *S. cerevisiae* is developed by Sonnleitner and Kappeli (1986), in which they present what is known as the respiratory bottleneck. The modelling assumption is that the microorganism has a maximal capacity of substrate consumption for biomass production, due to limitations in the respiration and when this capacity is reached ethanol is produced. The formulation of the model is rather simple and the results are in good agreement with experimental chemostat results. Their model describes the glucose consumption and the biomass, ethanol and CO_2 formation as well as oxygen consumption.

The key issue for process analysis is to model is the experimental fact that

ethanol is produced along with biomass if the glucose concentration is to high, this is what can be seen in a batch experiment starting with a high glucose concentration. The first phase of such an experiment is an exponential growth of the biomass, due to the auto catalytic nature of such process, with the following reaction scheme:

$$Glucose + Biomass \longrightarrow (1 + Y_{SX})Biomass \tag{2.1}$$

where Y_{SX} is the yield coefficient of biomass on glucose. At the same time the microorganism produces an overflow metabolite, in the case of yeast this metabolite is ethanol. The exponential phase ends when the glucose is consumed by the microorganism and it stops growing for a short period. After some time the microorganism starts to grow again with exponential growth rate, now utilising ethanol as substrate. The growth rate on ethanol is slower than for growth on glucose. The short intermediate phase where there is no growth the microorganism is adjusting to the new state of the environment and generating enzymes for utilization of ethanol for growth. This phase is called a lag phase. This is a special case where the lag phase occurs, the classical experiments are with two carbonhydrates, e.g. glucose and galactose where first the glucose is consumed and then the galactose after a lag phase.

Investigations of the uptake mechanism of the organism is studied by many researchers in order to break the lag phase, e.g. to obtain simultaneous utilization of different energy sources for higher efficiency, but this seems difficult due to the way the microorganism control reaction network. The control of uptake systems is described in section 2.3.2.

Another point is the utilization of other carbon hydrates than glucose. For a given microorganism there exists a range of carbon hydrates which can be metabolised however with different preferability. This means that the carbonhydrates are ordered such that they will be metabolised sequentially. One example is the glucose/ethanol pair discussed above. This can be illustrated by a batch experiment with three different carbon hydrates, first the most preferable one is metabolised then there is a lag phase then the second is metabolised and a new lag phase and then the last carbon hydrate is metabolised. Usually the ranking of carbon hydrates is related with the growth rate. How this phenomenon is regulated by the microorganism will be explained in more detail in section 2.3.1 and 2.3.2.

The organism has a number of uptake systems, some are very specialised towards one substrate, with high affinity but typically the capacity of such an uptake system is low. Other uptake systems have a high capacity but the affinity is lower and the substrate specificity is lower. This organisation of the uptake system makes sense, if there is very low concentration of substrate (energy source) it can pay off to synthesise the transporter enzyme for this substrate. If there was sufficient substrate there is no reason for the high affinity transporter since the transport through the low affinity transporter is high and therefore the synthesis is inhibited. From an energy efficiency point of view this make sense and could be the explanation for survival of these mechanisms in microbial evolution.

2.2.1 Operating Procedures

The surroundings of the microorganism is determining the behaviour of the cell. Therefore is interesting to see how the operation of different fermentation technology can change the surroundings, to obtain the desired behaviour.

Typically the production in the biochemical industry is carried out in a number of different operational environments such as pure batch, fed-batch or/and continuous operation. The difference between the three types of operation are being exploited.

The batch operation is being used for raising the concentration of the microorganism from an initially very low concentration to a reasonable level. The batch operation is characterised by mixing all substrates and microorganism.

The fed-batch operational mode can be used to increase the reactor volume to a desired level without having to do a batch in a very large volume, where the biomass concentration evidently will be very low at least initially. Another advantage of the fed-batch operation is that the substrate concentration in the reactor can be kept in control relatively low such that substrate inhibition and metabolic overflow can be avoided.

The continuous operation is used when steady state operation is desired. Continuous operation is to be understood in the sense of some hundred hours before shut down. The reason for not running longer can be many depending on the system and organism. The main reasons are risk of contamination of the process, instability of the genes/plasmids (the risk of spontaneous mutations) and the necessity of cleaning of the equipment due to wall growth.

2.3 Biochemistry - Basics

This section will describe basic biochemical phenomena that occur inside the living cell, which are important from an modelling prospective.

The number of components produced in a single microorganism is so large that the microorganism has to structure it in some way Lengeler (1996a). If the microorganism produce all components all the time the need of energy would be increased. One simple example is if there is an amino acid available in the surroundings the microorganism has mechanisms to detect it, have an uptake mechanism along with this the machinery for constructing the amino acid is shut down. The information is handled at many different levels in the individual cells. The total information in the genes defines the set of possibilities of the cell, whereas the environment determines the actual behaviour of the cell. For the cell to be able to read the stored genetic information it is converted into proteins, which has a specific property e.g. transport, reaction catalysis. The conversion of genetic information to proteins is carried out in two steps transcription and translation. Both these steps can be regulated by different mechanisms. To provide an understanding of these mechanisms a slightly more detailed description of the gene or more precisely the DNA will be given.

The basic information in the cell is contained in the genes which consists of

a double string of DNA (deoxyribonucleic acid), which is composed of four different bases, Guanine, Cytosine, Adenine and Thymine. The bases forming hydrogen bonds in pairs (Guanine-Cytosine) form 3 bonds and (Adenine-Thymine) form 2 bonds. Each of the strings of the DNA will then be a mirror image of the other string, meaning that if one string is ATTGCTTGGACCA the other string will be TAACGAACCTGGT, the two strings are matched together and form a double helix structure. The process of transcribing the DNA sequence into m-RNA is carried out by RNA polymerase. RNA polymerase consists of two main parts, a σ part and a core part, which consists of a complex of four polypeptides. The transcription mechanism may be described in four steps

1. Binding of the RNA polymerase to the DNA double string at the promotor site, forming a complex between the RNA polymerase and the σ part of the RNA polymerase.

2. The DNA double string opens up and is ready for transcription.

3. The σ part of the RNA polymerase leaves the core complex.

4. Transcription is initiated and polymerase moves along the DNA producing the mRNA, this process is called elongation since the RNA strand is elongated as the polymerase moves along.

The promotor site of the DNA is a specific sequence of amino acids which is recognised by the RNA polymerase. There are special sequences in special regions of the promotor that usually are observed, namely the -10 and -35 boxes. The specific sequence of these regions affects how strong the bonding of the polymerase to the DNA string is. Upstream from the promotor there is a region called activation site. At this site some molecules can bind to the DNA string and facilitates the binding of the RNA polymerase. Downstream from the promotor site there is a region called the operator. In this region molecules (repressor) bind to the DNA just downstream of the promotor and binding of the RNA polymerase is then not possible because the promotor is blocked by the repressor. Now we have what is necessary for controlling the transcription of DNA, an activation site for increasing the transcription (positive control) and a repressor for negative control. How these function is illustrated by the *lac* operon, in section 2.3.1.1.

2.3.1 Operons

In this and the following sections a description of the biochemical organisation inside the cell will be given based upon Lengeler (1996a). The described biochemical details is mainly relevant for bacteria but the overall mechanisms are general.

The way the microorganism keeps track of the immense number of chemical components that has to coordinated and structured is through control at differ-

ent levels, both positive and negative control are applied in order ensure the coordination. The reason for the necessity of the structuring is obvious.

- There is no need for synthesising e.g. amino acids if they are available.
- All the enzymes in the degradation of a substrate are needed simultaneously, otherwise a lot of energy will be wasted if enzymes are produced in unequal amounts.

To be able to control the expression of a number of proteins their genes are aligned in a series, a transcriptional unit. A transcriptional unit with a number of genes with common transcriptional control is called an operon. It is possible to control the expression of the whole operon by controlling the transcription of the unit. If the pathway is complex and consists of many enzymatic steps there might be good reasoning behind splitting the units. One reason is that the transcription tends to be unstable, in the sense that the polymerase jumps of with a certain frequency, this gives a maximal length of a operon. How the genetic information for such larger pathways are organised is discussed in section 2.3.2

2.3.1.1 The *lac* Operon

To describe the system of the *lac* Operon in *E-coli*, consider the following experiment. *E-coli* is grown on glucose and the genes in the *lac* Operon are not expressed at a significant level. The cells are then transferred to a medium with lactose. The first couple of minutes it appears as if nothing happens, growth has stopped. After a few minutes the cells starts growing again. During this lag phase the cells adjust to the new environment and synthesise the enzymes needed to grow on lactose. Details on the regulation of the *lac* operon are shown in figure 2.1. As long as glucose is present the *lac* Operon is repressed by the *lac* repressor. The *lac* repressor is a tetramer consisting of four identical polypeptides, which is allosteric, i.e. it can exist in two different conformations. In one conformation it binds to the operator of the operon and thereby blocks the transcription. In the other conformation it binds to allolactose, which is a derivate of lactose. Allolactose is in this context called an inducer, because it induces the transcription of the operon by releasing the operator and leaving the promotor free for the polymerase to initiate the transcription. This is not the whole story because if the experiment was carried out with a mixture of glucose and lactose, lactose would not be consumed before the glucose was depleted. This make sense since *E-coli* can metabolise glucose more easily than lactose. The cell can turn on the operon when glucose is gone in the following manner. When glucose is consumed the level of a small nucleotide, cyclic-AMP (cAMP), increases as an indication of energy starvation. The operon is activated by a complex of a protein, CRP (catabolite repression protein) and cAMP. This complex binds to a region upstream from the promotor and facilitates the binding of the polymerase.

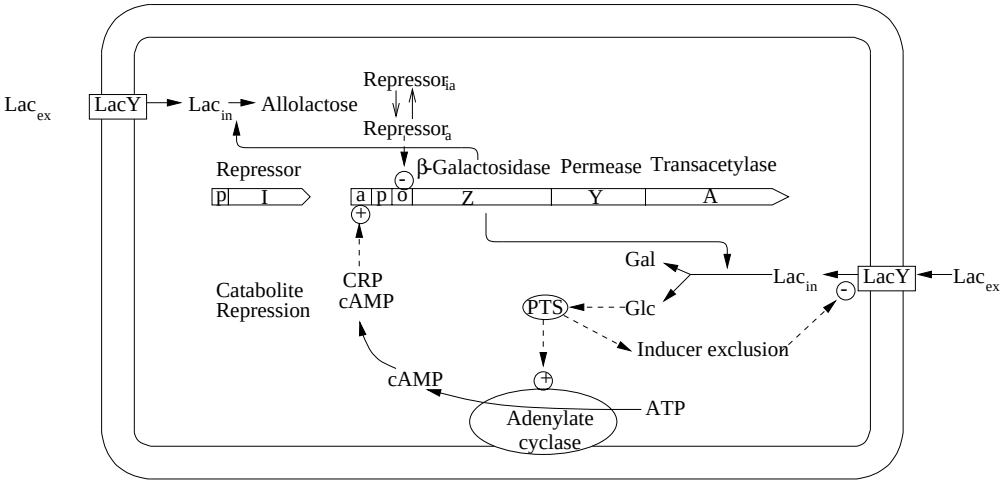

Figure 2.1: Operation of the Lac Operon. The control of transcription is not perfect and therefore there will always be small amount of enzymes present. In presence of glucose the cAMP level is low and the activation site (a) is free and binding of polymerase to the promotor (p) is not possible. If glucose is absent the cAMP forms a complex with CRP and binds to a. If lactose is unavailable the repressor (i) is in its active state and blocs the transcription, since it is bound to the operator site (o). In the presence of lactose the repressor is inactivated by allolactose which is produced from lactose by β galactosidase. Now the transcription of LacZ (β galactosidase), LacY (Permease) and LacA (Transacetylase). If glucose is added, growth on lactose is stopped through deactivation of the LacY by inducer exclusion which is related to the PTS system, see Lengeler (1996b)

Another way of looking at the lac operon is as a sensory system for lactose in the absence of glucose. When glucose is unavailable for the organism the CRP-cAMP complex activates the promotor region. Only when lactose is present repression at the operator site is released and transcription will run at full strength. This way of controlling transcription can be seen as optimal use of resources, only the necessary enzymes will be produced.

2.3.2 Modulon Modelling

Some of the pathways in living cells are significantly more complex than the lactose pathway. The requirement for coordinated expression of enzymes is still evident, but there is a limitation on the length of a transcriptional unit. The limitation in the length of a transcriptional unit is determined by the binding of the polymerase on the DNA. The polymerase will with some frequency drop off the DNA and the rest of the gene will not be transcribed. If a transcriptional unit consists of 10 genes and the polymerase drops off around gene 6-8 the last couple of genes will most likely never be transcribed, since transcription has to be initiated from the promotor site. To overcome these limitations the

genes of larger pathways are split up into a number of operons with a common regulatory system. By having the same regulatory system the expression of the genes in all operons can be coordinated and new detailed regulation can be introduced, e.g. expression of genes suitable for anaerobic/aerobic conditions can be controlled. Such a collection of operons is called a regulon but is basically the same as an operon.

On top of the operon/regulon layer of regulation there are systems which keep track of the need for energy, which amino acids should be synthesised etc. These systems are of more global nature from the cell point of view. The cAMP-CRP regulation of the Lac operon is in fact a global regulator which activates most of the catabolic operons/regulons in the same way as the Lac operon, figure 2.2. This makes sense since the cAMP-CRP complex is an indication of starvation. As mentioned in section 2.3.1 the cAMP level increases when the glucose flux was decreasing. This is true for a number of easy metabolisable substrates.

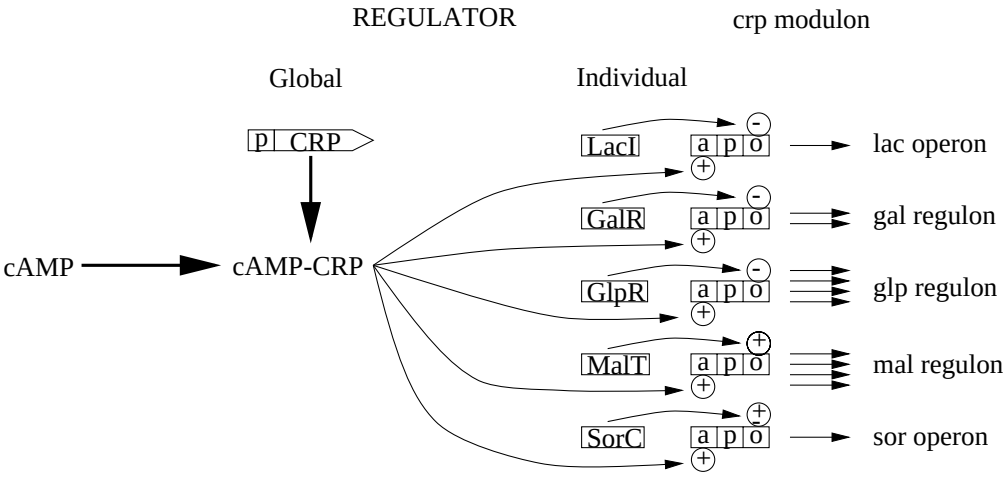

Figure 2.2: The function of part of the CRP modulon, lac : lactose, gal : galactose, glp : glycerol, mal : maltose, sor : sorbose

Most metabolic enzymes are controlled is this way at the transcriptional level, and such control of global properties is called a regulon.

2.3.3 Control of reactions

There is another level of control in the microorganisms which was touched upon in the section of the *lac* operon. The control of the individual reactions. It is a fact that most (if not all) reactions in a living organism are catalysed by an enzyme. In order to fine tune the rate of reaction of a pathway the enzymes can be modified by other molecules such that the characteristics change. The change can either be an increased or a decreased reaction rate. The enzymes are large macromolecules which in some cases consist of a number of subunits which interact. It is the location of binding sites in the macromolecules that

determine the affinity for a specific component. The affinity can be changed by binding of other molecules which will cause a change in the folding of the macromolecule and thereby a change in the relative location of binding sites. Such a change of folding is called allosteric control. Allosteric means a change of shape and function by binding of a small molecule Weaver and Hedrick (1992). As in the *lac* operon example the inactivation of the repressor protein is by allosteric control. The properties of the allosteric protein is that it has two binding sites one for the control molecule and one for the reactant. The typical reaction times for this type of control mechanism is in the range below a second, whereas the control at the gene level is in the range of minutes. The reason for this time gap is clear, since the release of the control molecule is only one reaction, whereas the time to transcribe and translate the gene into the particular enzyme, which involves some hundred reaction steps, will necessarily be slower.

2.4 Unstructured modelling

In this and the next section the ideas in modelling the behaviour of microorganisms are discussed. When modelling a microorganism there are many different ways to do it, the approach is mainly determined by the modelling purpose as mentioned in section 2.1. By unstructured models is meant that the biomass is unstructured (considered as one component) and the fermentation process is normally modelled by one or a few reactions only.

To model the overall conversion of substrate into biomass and other products, one can use a fairly simple model like the one given below for a continuous operating tank reactor, assuming unstructured biomass, using dynamic mass balances for biomass, X, and substrate, S

$$\frac{dX}{dt} = Y_{SX}\mu X - DX \tag{2.2}$$

$$\frac{dS}{dt} = -\mu X + D(S_f - S) \tag{2.3}$$

μ is the reaction rate or growth rate of the biomass, as mentioned Y_{SX} is the yield coefficient, D is the dilution rate given by $D = F/V$, where F is the feed rate to the reactor, V is the reactor operating volume and S_f is the substrate feed concentration. μ is typically modelled as Monod kinetics,

$$\mu = \frac{\mu_{max} S}{K + S} \tag{2.4}$$

μ_{max} is the maximal growth rate, and K is the affinity constant. The simplicity of the model is amazing seen from the point that a cell is a complicated living organism. The modelling is looking at the microorganism as unified, only converting substrate into biomass according to equation 2.1. The conversion of substrate into some by products as ethanol described earlier can easily be

added to the model structure of 2.2 and 2.3. There are some limitations to this modelling approach simply if the interest is the dynamics of the intracellular components. Another limitation which is generally recognized as a flaw of unstructured models is that they are not able to predict transient behaviour of the process. This is also to be expected since the largest changes in the intracellular compositions of an organism are occurring during transients, e.g shift from one substrate to another or step up in dilution rate. One way to circumvent the problem of describing the dynamical properties of the process have been to assign different parameter values for different types of operation. This approach will result in limited applicability e.g. when the model is to be used for control design and other application areas. For simple simulation studies the approach may be reasonable, but for increasing knowledge of the organism a structural modelling approach will be preferable, since the focus is on the cell metabolism. The approach is different as the unstructured models will mainly be applicable for continuous operation since they represent transient behaviour rather poorly. Structured models tend to be well suited for applications in batch and fed-batch operation since these models better can describe transient/changes in the composition of the biomass.

2.5 Structural modelling

To overcome the limitations of the unstructured modelling approach briefly touched upon in the previous section the biomass is divided into components which hold some key properties. This is done to be able to describe in more detail and with higher precision some key features of the microorganism behaviour. The approach is very useful to get structure into the modelling and to get a structure of the organism into the model. By structure of the organism is meant both the physiological structure (like filamentous fungi) and the biochemical structure of the reaction network in the cells. The structure of the reaction network is the interplay between the genes, proteins and the different substrates and how their levels influence the total network behaviour. The classical way to structure the biomass is to divide it into a structural and an active part. The structural part is cell wall and other components which play no active part in the degradation of substrates or in the generation of products. The active part of the biomass is defined Nielsen and Villadsen (1992), as the proteins, enzymes, precursors and active metabolites.

2.5.1 Metabolic network

In this section and the following the ideas of metabolic flux analysis and metabolic control analysis is introduce. These two areas have been developed over the last couple of decades, starting with Kacser and Burns (1973), and are helpful tools in the model building for microorganisms. These ideas are then used in a very promising way to optimise the flux of given parts of the metabolism. A metabolic network is defined as the sequence of reactions which describe

the degradation of metabolites (catabolism) and building of the intracellular components (anabolism) of the cell. The network of the cell reactions are well described in the literature in terms of the basic reactions which forms the backbone of the reaction network.

2.5.2 Metabolic Flux Analysis

Metabolic flux analysis is an analytical tool for determining the intracellular fluxes through a specified metabolic network. The basis for the flux analysis is the knowledge of the reactions in the network under consideration.

When the reactions are defined their stoichiometry provide relations between the chemical components. These relations are used in the calculation of the fluxes in the network and to reduce the need for measurement in the analysis. The mathematical setting will be described in the next section. It can be used for determination of which pathways are active and which are not, by assuming different metabolic network structure, at given conditions.

2.5.2.1 Mathematical setting

The field of flux analysis is well described in the literature and textbooks such as Nielsen and Villadsen (1994). Flux analysis is by nature a simple linear steady state modelling approach which will result in a measure of the internal fluxes in the organism at a given steady state. Since the concern is steady state the relevant type of operation is chemostat at fixed dilution rate. Numerous examples are given in both textbooks and in the literature.

The basis of the flux analysis is the stoichiometry of the involved reaction network. The stoichiometry is set up in a matrix notation

$$\begin{bmatrix} ABC \end{bmatrix} \begin{bmatrix} s \\ p \\ X \end{bmatrix} = T \begin{bmatrix} s \\ p \\ X \end{bmatrix} = 0 \tag{2.5}$$

Where the sub matrices A, B, C contains the stoichiometric coefficients for the reactions in the network. The concentrations s, p, X is substrates, products and internal components respectively. The volumetric reaction rates q of the reaction in 2.5 are given by

$$q = T^T r x \tag{2.6}$$

where x is the biomass concentration. The rate vector r is the unknown and will be split up into a measurable part and a calculated part $r = \begin{bmatrix} r_m & r_c \end{bmatrix}^T$, it is further assumed that there is a number of pseudo steady state intermediates in the organism, this is what helps in the calculation. The result is that equation 2.6 can be rewritten into

$$\begin{bmatrix} q_m \\ q_c \\ 0 \end{bmatrix} = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \\ T_5 & T_6 \end{bmatrix} \begin{bmatrix} r_m x \\ r_c x \end{bmatrix} \tag{2.7}$$

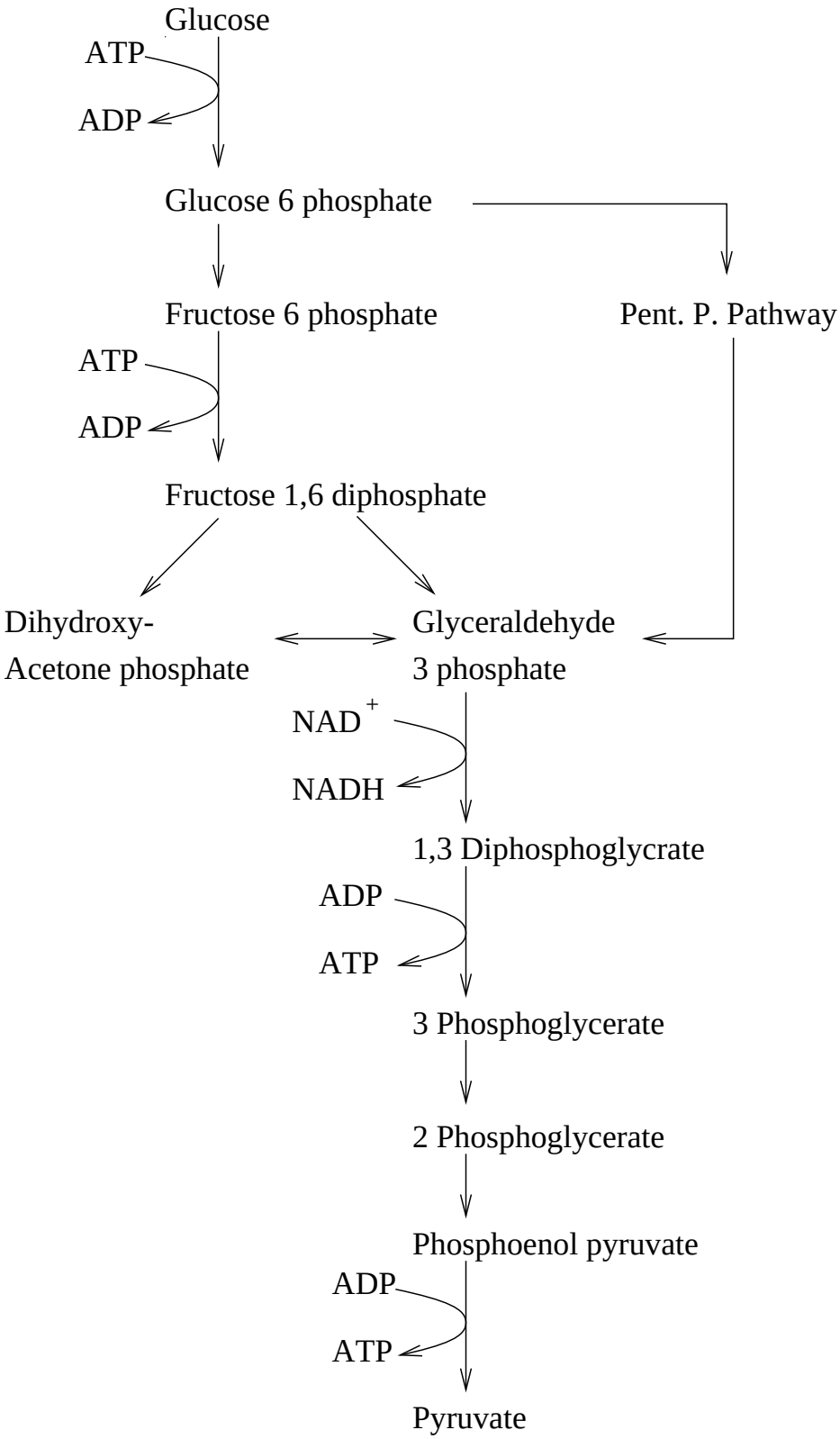

Figure 2.3: Embden Meyerhof Paranas (EMP] Pathway and Pentose Phosphate pathway, the backbone of glycolysis.

The solution of this set of linear equation is given by Nielsen and Villadsen (1994):

$$T_7 = T_1 - T_2 T_6^{-1} T_5 \tag{2.8}$$

$$r_m x = T_7^{-1} q_m \tag{2.9}$$

$$r_c x = -T_6^{-1} T_5 T_7^{-1} q_m \tag{2.10}$$

$$q_c = (T_3 - T_4 T_6^{-1} T_5) T_7^{-1} q_m \tag{2.11}$$

The above formulation is for simple calculation where the size of the problem is low and therefore solution of the problem by hand is an option. For practical application the formulation in equation 2.12 is simpler from an implementation point of view, but the result is clearly the same.

$$\begin{bmatrix} r_m x \\ r_c x \end{bmatrix} = \begin{bmatrix} T_1 & T_2 \\ T_5 & T_6 \end{bmatrix}^{-1} \begin{bmatrix} q_m \\ 0 \end{bmatrix} \tag{2.12}$$

$$q_c = \begin{bmatrix} T_3 & T_4 \end{bmatrix} \begin{bmatrix} r_m x \\ r_c x \end{bmatrix} \tag{2.13}$$

Either way the measurement of volumetric net formation rate of the components in q_m, one can calculate the fluxes through the cellular reactions and the remaining volumetric rates. Care should be put into the analysis of which measurements to use in the calculation. Ideally the system should be overdetermined to check for measurement errors. The main point is the properties of the T matrix in equation 2.12 which is affected by the elements of q_m.

Since the background for the analysis is the stoichiometry of the reactions, as given in equation 2.5, simple linear reaction steps are included as overall reactions. The reason for this is that linear combinations of the intermediate reactions will result in the overall expression, thus rank problems of the T matrix will arise in the solution of the flux analysis problems. The result is that only branch points are included as the intermediates in X.

2.5.3 Metabolic Control Analysis

Metabolic control analysis is used to quantify how the fluxes of a pathway are influenced by the kinetics and control imposed on the reactions by other components. The investigation will show which steps are rate determining for the given pathway. The control analysis can be applied to both linear reaction sequences as well as branched pathways.

The components that will influence the reaction rate could be the enzymes, that catalyse the reactions, and other chemical components which influences the reaction rates by inhibition/activation. The result of the control analysis is control coefficients, which indicates where in the network the benefit of a change will have the greatest impact on the flux. The idea behind this concept originates back to Kacser and Burns (1973). The set of reactions is L enzymatic reactions from a substrate to a product, with $L-1$ intermediates. The basis is

the definitions of *flux control coefficients* by

$$C_i^{J_j} = \frac{E_i}{J_j}\frac{\partial J_j}{\partial E_i} \quad i = 1,\ldots,L \tag{2.14}$$

Where the J_j is the flux through the j reaction and E_i is the enzyme activity of the i'th enzyme. The Flux control coefficient can be viewed as a measure of the individual enzymes influence on the total flux through the pathway.

Similar definitions exists of *concentration control coefficients.*

$$C_i^{X_j} = \frac{E_i}{c_j}\frac{\partial c_j}{\partial E_i} \quad i = 1,\ldots,L \tag{2.15}$$

The concentration control coefficients is a measure of the influence of the enzyme activities on the level of the metabolites in the pathway. Another element in the metabolic control analysis is the *elasticity coefficient*

$$\varepsilon_{X_j}^i = \frac{c_j}{v_i}\frac{\partial v_i}{\partial c_j} \quad i = 1,\ldots,L \tag{2.16}$$

The elasticity coefficients is a measure of how the levels of metabolites influence the rate of the enzymatic reactions. Some simple relations is imposed by the structure of the coefficients

$$\sum_{i=1}^{L} C_i^{J_j} = 1 \quad j = 1,\ldots,L \tag{2.17}$$

$$\sum_{i=1}^{L} C_i^{X_j} = 0 \quad j = 1,\ldots,L-1 \tag{2.18}$$

$$\sum_{i=1}^{L} C_i^{J_j}\varepsilon_{X_j}^i = 0 \quad j = 1,\ldots,L-1 \tag{2.19}$$

$$\sum_{i=1}^{L} C_i^{X_j}\varepsilon_{X_j}^i = -1 \quad j = 1,\ldots,L-1 \tag{2.20}$$

$$\sum_{i=1}^{L} C_i^{X_k}\varepsilon_{X_j}^i = 0 \quad j = 1,\ldots,L-1 \quad \text{and} \quad k \neq j \tag{2.21}$$

The method is intended to point out where to make modifications by genetic engineering. The main problem is to determine the flux control coefficients since the intracellular measurement is hard to get, except through NMR studies of labeled components. By combining the summation rules the following equation will result

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1L} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{L-11} & \varepsilon_{L-12} & \cdots & \varepsilon_{L1L} \end{bmatrix} \begin{bmatrix} C_1^{J_1} & -C_1^{X_1} & \cdots & -C_1^{X_L-1} \\ C_2^{J_1} & -C_2^{X_1} & \cdots & -C_2^{X_L-1} \\ \vdots & \vdots & \ddots & \vdots \\ C_L^{J_1} & -C_L^{X_1} & \cdots & -C_L^{X_L-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.22}$$

The solution of the equation in a compact notation is

$$EC = I \Rightarrow C = E^{-1} \tag{2.23}$$

Where the first column of C contains the flux control coefficients needed for the analysis of the pathway. The idea is that all the enzymes in the pathway should have an even control on the flux through the pathway. Therefore the target point for the L control coefficients is a value close to $1/L$. If in some example there is a large coefficient that particular enzyme will have a large impact in the flux compared to the other enzymes. The method should be used as a guide for the genetic modifications of the organism since the modified organism will not necessarily show an increased flux but a more evenly distribution of the flux control. Verifications of the results by modifying the genes of the organisms are often disappointing due to the complexity of the cells control structure.

2.6 Optimized Network

By defining the metabolic network it is also possible to apply an optimization scheme to the model by setting a number of parameters free for the optimization routine to determine. This work was initiated by Baily and co-workers Hatzimanikatis et al. (1996) and gives a qualified guess for target points to be addressed by genetics. The approach received only little attention in the literature e.g. Hatzimanikatis et al. (1998). By formulation of a power law model for the kinetics and using a logarithmic transformation it is possible to formulate mixed integer linear programming problem. The model should contain all the information available on the interactions in the reaction network. Meaning that information on enzyme activities (levels) as functions of substrates and other metabolites as well as information on the influence on the reaction. The attractive thing about the model optimization is that the solution of the linear programming problem is simple in terms of mathematics. The approach has to include reasonable bounds on the free variables and care should be taken since the real parameters are hidden due to the transformation. Since the models are predefined the applicability is in some cases limited but one should then consider the alternative which would be nonlinear modelling and nonlinear programming which is by no means simple and straight forward from a solution point of view. Solution of nonlinear programming problems will typical involve linearisation and numerous iterations. This aspect will be discussed in more detail in chapter 6

2.7 Summary

In this chapter an introduction to the following chapters have been provided and the point of view on the cell as a highly organised unit, not just a bag of enzymes, is presented. The main idea with this chapter is to guide readers which are unfamiliar with the terminology in the area of biotechnology into this area.

The discussion on the behaviour of microorganism at different conditions lead to the conclusion that understanding of the biochemistry is necessary. In section 2.3 the control of enzyme expression is illustrated through the lac operon and the CRP modulon. With this in mind it is evident that unstructured models will fail to describe transient phenomena since the internal control of enzymes is dominant at these conditions. Therefore structured models are preferable for description of the behaviour such transients. Finally flux and control analysis are presented and an application of optimization of metabolic control is briefly touched upon in the last two sections of the chapter.

3

Modelling of Yeast Cultivations

The modelling of microorganisms have received increasing interest from researchers with different background. To benefit form the different inputs strong cooperation between engineers and biologists is necessary. Depending on the purpose of the model, different levels of detail are required in the modelling. Here is presented a model for investigation of the behaviour of yeast around the critical dilution rate where the microorganism may shift metabolic pathway.

3.1 Introduction

The modelling of yeast cultivation have been investigated by many researchers and for many different purposes. The approach taken by biologists is quite different from the approach taken by an engineer due to their different background and points of view. All looks at some specific detail of the organism and its behaviour and model it in some narrow operation region.

From a production point of view the interesting point is where the yield of biomass is maximal, when the production of special products typically is growth related.

The model presented here constitute the background for the analysis of yeast cultivation in the subsequent chapters on bifurcation analysis and control design.

The reason for development of models in the area of biotechnology is the increased use of these processes in the industry. The amount of knowledge in this field is increasing, and the understanding of the underlying biochemical mechanisms is beginning to emerge in a way that will facilitate the mathematical modelling based on engineering principles. Therefore the development of models in the area of biochemical engineering is an essential prerequisite for the analysis techniques which are being applied to more traditional chemical engineering processes.

Based on the knowledge of the metabolic pathway a formulation of key components is posed. There have been some evidence in the literature that there are operational problems around the critical dilution rate where there is maximal production of biomass, see Rieger et al. (1983). Therefore the task is to develop a model of what occurs when yeast is shifting metabolism. As discussed in section 3.2 the growth rate of biomass is limited, but it is observed that the

substrate uptake rate does not have the same limitation. Therefore the substrate must be directed in some other direction, in yeast the main byproduct is ethanol. This shift in the metabolism is from oxidative to oxido reductive conditions, where the substrate is only partly oxidised. The chapter will describe modelling of the mechanism which occur around the shift in metabolism.

The shift occurs when the organism is growing and the flux of the substrate to biomass is maximal. If the substrate flux at this point is increased the biomass flux will not increase. The organism has reached its maximal capacity, in this case the oxidative capacity. In other words the cell will have to cope with the excess of reduction equivalents. The presented model have been developed over the years at the department, Rotbøll (1992).

3.2 Model developments

The focus of the modelling study is to describe the metabolism around the critical dilution rate, the shift in the metabolism was traditionally known as the Crabtree effect, i.e. an inhibition of the oxidative system by high glucose concentrations. Now it is generally accepted that the formation of ethanol at aerobic conditions is a consequence of limitations in the oxidation of pyruvate, e.g. in the respiratory system Alexander and Jefferies (1990). The shift in the metabolism has been modelled by many people, the most widely known and cited model is the unstructured one by Sonnleitner and Kappeli (1986), where they introduce the modelling concept of a bottleneck in the metabolism, which is illustrated in figure 3.1.

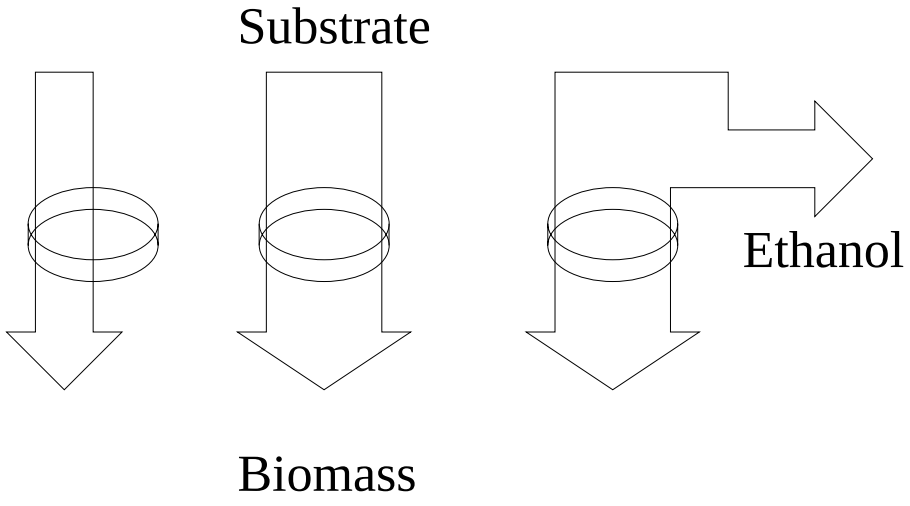

Figure 3.1: Bottleneck modelling concept for microorganism, based on the assumption that there is a limitation in the capacity for substrate utilization for biomass production, surplus of substrate will be directed towards ethanol

The figure shows that at low flux of substrate is all directed towards biomass, up to a certain limit, then the biomass flux is maximal and if this limit is

exceeded the excess substrate flux is directed into ethanol. In the model of Sonnleitner and Kappeli (1986), which is an unstructured model of substrate, biomass, ethanol and oxygen. The implementation of the switch is modeled by an IF THEN statement. The reason for this approach is that there is not sufficient information in the model to describe the underlying biochemistry, one could say that the model is too simple to describe the real shift. The simplicity of the model is appealing since it can describe continuous cultivations in reasonable agreement with experimental results. The drawback of the simplicity is that the understanding of what happens around the switch inside the cell is not included in the model. The unstructured model is only able to describe one type of experiments e.g. batch, fed-batch or continuous operation on a single substrate, with on set of parameters. If another experiment should be explained by the same model then the parameters must be changed, e.g. Sweere et al. (1988).

The way detailed models are constructed is by dividing the biotic phase, which encompasses what is happening inside the cell, into different compartments, typically the first division is into a compartment with all growth related active enzymes and one with all the DNA, different macromolecules, cell wall components including transport enzymes. Examples of such models can be found in the review by Nielsen and Villadsen (1992). They propose a structured yeast model where the shift in metabolism is based on the ATP balance. Their model serves as a basis for the model described here, however the mechanism of the shift is different.

For an example of a highly detailed model see Steinmeyer and Shuler (1989), but the problem with this level of detail is that the number of parameters is increasing and the estimation of these parameters become troublesome and a lot of detailed experiments with many measurements are needed for reliable parameter estimation. The model described below is based on the observation that there is accumulation of both pyruvate and acetaldehyde around the critical dilution rate, see Postma et al. (1989).

3.3 The model background

The modelling of a biotechnical fermentation process is based on many simplifying assumptions. The assumptions are divided into those concerning the fermentor/medium, (the abiotic phase), assumptions concerning the microorganism, (the biotic phase), and structural assumptions.

Abiotic assumptions:

1. The reactor is ideally mixed, this might be true for small scale lab fermentors but certainly not for industrial scale. The experiments used for parameter estimation and verification are all obtained from labscale.

2. Other essential growth components, such as temperature, pH, oxygen and nitrogen sources, are maintained at their target values, and therefore their influence the growth rate can be neglected. If the first assumption

holds then this is more realistic, but high cell densities might complicate the matter of ideally mixed.

3. The model is only concerned with the growth on glucose and ethanol as limiting substrates. This might reflect the conditions in the lab but not complex substrate mixtures often used in industry.

Biotic assumptions

1. The cell population is assumed to be homogeneous. This is a reasonable approximation, but one should be aware of the fact that chemostats have been reported to behave in a cyclic manner, due to synchronization of the cell cycle over some range of dilution rates. This will not be possible to see with this model.

2. The biotic phase is assumed to be uniform, i.e. no gradients inside the cells (well mixed cytoplasma).

3. All components which are not included in the model are assumed to be in pseudo steady state.

4. The composition of the biomass is independent of the fermentation conditions and constant.

5. Stoichiometric coefficients are constant.

6. Intracellular glucose concentration can be decoupled by the extracellular glucose concentration.

7. Pyruvate is distributed according to the difference between the intra- and extracellular pH and the intracellular pyruvate concentration can therefore be decoupled with the extracellular pyruvate concentration.

8. Ethanol, acetaldehyde, CO_2 and dissolved oxygen diffuses rapidly through the cell and the intra- and extracellular concentrations are equal.

Structural assumptions

1. The selected components are assumed to describe the basic reaction network. Since there is good agreement with experimental result in literature this is reasonable. It should be mentioned that the model structure is a simplification of the real structure.

2. The signal which controls the shift in the metabolism is assumed to be closely related to acetaldehyde. This assumption is supported by some experimental results. This component is central in the structure of the reaction network.

3.4 Reaction Network

The model is extended from that of Nielsen and Villadsen (1992). Two new metabolites are introduced in the cell, pyruvate and acetaldehyde, two enzymes are also included in the model to help the reasoning and the modelling. The development is done to overcome the artifact in the models of both Sonnleitner and Kappeli (1986) and Nielsen and Villadsen (1992) where a simple IF THEN type of statement is included to model the limitation in the capacity. The physiology of the microorganism do not have an IF THEN type of functionality, but as indicated in the chapter on biochemistry, switches in reaction network consists of allosteric enzymes. The Control of transcription/translation and regulation of reaction rates will result in the IF THEN kind of behaviour. The pathways used in this model are illustrated in figure 3.2

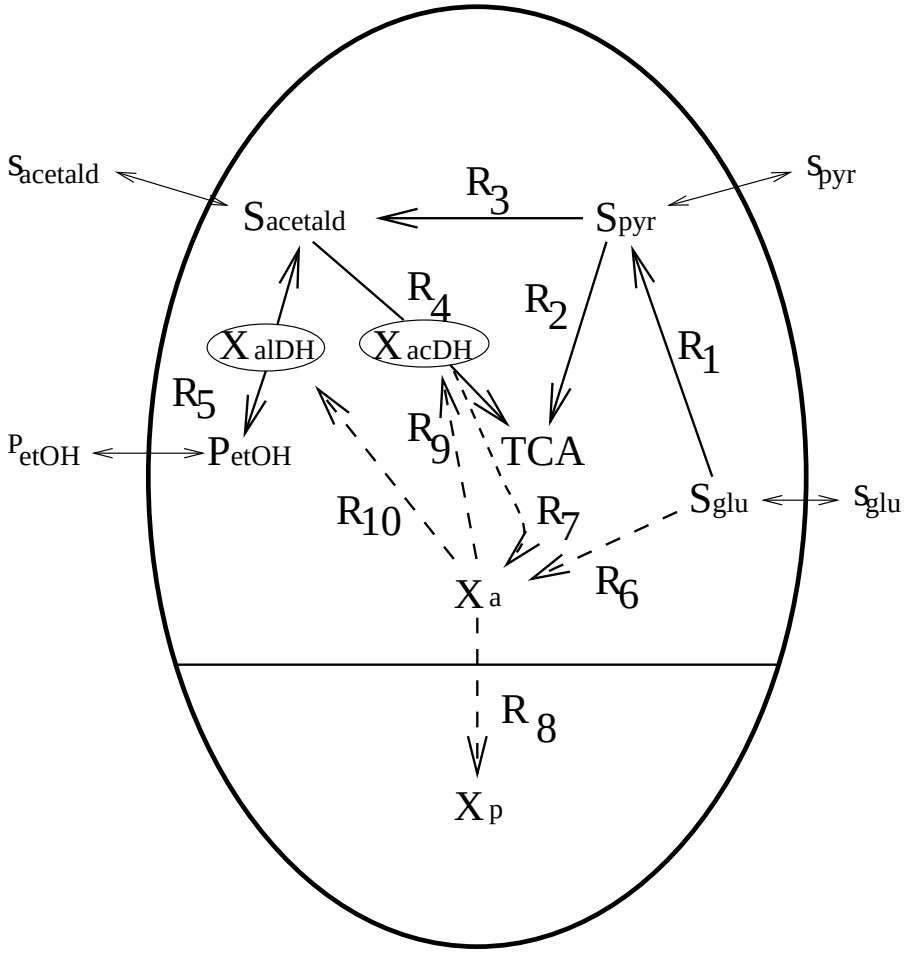

Figure 3.2: Structure of the yeast model

3.4.1 Stoichiometry of reactions.

In this section the reactions described in the previous section is formulated and stated. The kinetic parameters are not estimated in this work but simply taken from the work of Rotbøll (1992) and Faxø (1995). The basis of the stoichiometry is given on a g/g basis in table 3.1. The basis for the stoichiometry can be found in many textbooks on biochemistry. The structure of the reactions are shown in figure 3.2. Only the modelled components are included in the reactions, so the balancing of elemental composition is not fulfilled.

R_1	S_{glu}	$\rightarrow$	$0.9778S_{pyr}$
R_2	$S_{pyr} + 1.0909O_2$	$\rightarrow$	$1.5CO_2$
R_3	S_{pyr}	$\rightarrow$	$0.5S_{acerald} + 0.5CO_2$
R_4	$S_{acerald} + 2.1818O_2$	$\rightarrow$	$2CO_2$
R_5	$S_{acerald}$	$\rightleftarrows$	$1.045P_{EtOH}$
R_6	$S_{glu} + 0.063O_2$	$\rightarrow$	$0.732X_a + 0.127CO_2$
R_7	$S_{acerald} + 1.0178O_2$	$\rightarrow$	$0.850X_a + 0.4431CO_2$
R_8	X_a	$\rightarrow$	X_p
R_9	X_a	$\rightarrow$	X_{acDH}
R_{10}	X_a	$\rightarrow$	X_{alDH}

Table 3.1: Stoichiometry of reactions in the metabolic network, g/g basis, for the reactions in figure 3.2

3.4.2 Reaction kinetic expressions

The kinetics of the individual reactions is the most unsure area of the modelling, therefore great care in the description of the reaction kinetics is needed. The individual reactions are all described below and their rate expression shown in equation 3.1-3.10.

R_1 The reaction rate from glucose to pyruvate is a simplification of the glycolysis, since this is a result of a sequence of reactions which here are lumped into one. The reason for this lumping of reactions is that the reaction sequence is controlled at all steps by enzymes, and therefore the overall rate can be modeled as one effective reaction. The rate expression consists of two terms, the first term describes the Michaelis-Menten type kinetics. The second term is activated by acetaldehyde and is therefore only active when ethanol is being produced. The term is empirical and can be interpreted as a consequence of the saturation of the energy producing system.

R_2 Pyruvate can either enter the TCA cycle in the mitochondria where it is oxidised or be converted to acetaldehyde. The first reaction is repressed by glucose, Postma et al. (1989). That effect was not sufficient and therefore the extra inhibition term by acetaldehyde is introduced.

R_3 The reaction rate of pyruvate to acetaldehyde is described by simple Michaelis-Menten type kinetics.

R_4 Acetaldehyde can either be converted into ethanol by alcohol dehydrogenase or be used in the TCA cycle and end up as carbon dioxide. The latter is modelled with the enzyme acetaldehyde dehydrogenase and is actually split up into two parts. The reaction rate through acetaldehyde dehydrogenase is modelled by simple Michaelis-Menten kinetics with subtraction of the amount that is directed into biomass. This flux is described in reaction R_7.

R_7 The reaction from acetaldehyde into biomass occurs only when the cells are grown on ethanol and glucogenesis occurs. Glucogenesis is reverse glycolysis on the reaction level, but the enzymes are different and it is only switched on when glucose is depleted, actually only when all primary energy sources are depleted, but here glucose is only considered therefore the inhibition term on glucose.

R_5 The reaction to ethanol is reversible and the equilibrium is controlled by ethanol and acetaldehyde the rate expression is empirically determined.

R_6 The formation of biomass from glucose is described as a simple enzymatic reaction with Michaelis-Menten reaction kinetics, the reaction is a number of anabolic reactions pooled together into one, but as mentioned earlier this yield a reasonable approximation.

R_7 The reaction from acetaldehyde to biomass (the alternative pathway for biomass formation) is also described with Michaelis-Menten kinetics, but since it is only active when the growth is on ethanol, the term is inhibited by glucose.

R_8 The formation of the passive compartment is depending on the source of energy and is divided into two separate Michaelis-Menten kinetic expressions.

The two enzymes are formed from the active compartment.

R_9 Acetaldehyde dehydrogenase rate is formulated as two Michaelis-Menten expressions repressed by glucose, and fitted to the data in Postma et al. (1989).

R_{10} The Alcohol dehydrogenase rate is described by two terms, a first order term with respect to the active compartment and repressed by glucose to ensure a decrease with increasing glucose concentration below the critical bottleneck rate as observed by Postma et al. (1989). The second term is described as Michaelis-Menten kinetics with respect to acetaldehyde.

All the parameters are either taken from the literature or estimated by using experimental data from the literature. For more details on the procedure of the

parameter estimation see Rotbøll (1992). Lei et al. (1999) describe a modified version of the presented model, but the parameter estimation is improved to include both chemostat and batch data in the estimation.

The kinetic expressions is given in equations 3.1-3.10, the basis for the kinetic expressions is Michaelis-Menten type of kinetic which describes enzymatic catalysed reactions and as indicated in the text above there are some inhibitions which are included in the model to reflect the experimental observations. These inhibitions are somewhat empirical, but reflects the experiments rather well.

$$R_1 = \frac{k_1 S_{glu}}{S_{glu} + S_1} X_a + \frac{k_{e1} S_{glu} S_{acetald}}{S_{glu}\left(1 + K_{i1} S_{acetald}\right) + S_{e1}} X_a \tag{3.1}$$

$$R_2 = \frac{k_2 S_{pyr}}{S_{pyr}\left(M_2 S_{acetald} + 1\right) + S_2} \frac{1}{\left(M_{e2} S_{glu} + 1\right)} X_a \tag{3.2}$$

$$R_3 = \frac{k_3 S_{pyr}}{S_{pyr} + S_3} X_a \tag{3.3}$$

$$R_4 = \frac{k_4 S_{acetald}\left(k_{e4} - \frac{1}{\left(k_{i4} S_{glu} + 1\right)}\right)}{\left(S_{acetald} + S_4\right)} X_{acDH} X_a \tag{3.4}$$

$$R_5 = \frac{k_5\left(S_{acetald} - \frac{k_{r5} P_{EtOH}}{1 + k_{i5} S_{acetald}}\right)}{\left(S_{acetald} + S_5 + S_{e5} P_{EtOH}\right)} X_{alDH} X_a \tag{3.5}$$

$$R_6 = \frac{k_6 S_{glu}}{S_{glu} + S_6} X_a \tag{3.6}$$

$$R_7 = \frac{k_4 S_{acetald}}{\left(S_{acetald} + S_4\right)\left(k_{i4} S_{glu} + 1\right)} X_{acDH} X_a \tag{3.7}$$

$$R_8 = \frac{k_8 S_{glu}}{S_{glu} + S_8} X_a + \frac{k_{e8} P_{EtOH}}{P_{EtOH} + S_{e8}} X_a \tag{3.8}$$

$$R_9 = \frac{k_9\left(\frac{S_{glu}}{S_{glu} + S_9} + \frac{P_{EtOH}}{P_{EtOH} + S_{e9}}\right)}{\left(k_{i9} {S_{glu}}^2 + 1\right)} X_a \tag{3.9}$$

$$R_{10} = \frac{k_{10}}{k_{i10} {S_{glu}}^2 + 1} X_a + \frac{k_{e10} S_{acetald}}{S_{acetald} + S_{e10}} X_a \tag{3.10}$$

3.4.3 Macroscopic mass balances

The development of the mass balances follow straightforward from this point, with the reaction network as given in figure 3.2 and the kinetics given in equations 3.1-3.10 and the stoichiometry of the reactions from table 3.1. The assumptions of setting up the mass balances have been discussed in the previous sections. The different modes of operations are included through the dilution rate, D, which is either zero, varying in time or constant. D is defined as the volumetric flow rate divided by the actual operation volume, $D = F/V$.

$$\frac{dS_{glu}}{dt} = -(R_1 + R_6)X + (S_{glu,\,feed} - S_{glu})D \tag{3.11}$$

$$\frac{dS_{pyr}}{dt} = (0.9778R_1 - R_2 - R_3)X - S_{pyr}D \tag{3.12}$$

$$\frac{dS_{acetald}}{dt} = (0.5R_3 - R_4 - R_5 - R_7)X - S_{acetald}D \tag{3.13}$$

$$\frac{dP_{EtOH}}{dt} = (1.045R_5)X - P_{EtOH}D \tag{3.14}$$

$$\frac{dX}{dt} = (0.732R_6 + 0.850R_7)X - XD \tag{3.15}$$

$$\frac{dX_a}{dt} = 0.732R_6 + 0.850R_7 - R_8 - R_9 - R_{10} - (0.732R_6 + 0.850R_7)X_a \tag{3.16}$$

$$\frac{dX_p}{dt} = R_8 - (0.732R_6 + 0.850R_7)X_p \tag{3.17}$$

$$\frac{dX_{acDH}}{dt} = R_9 - (0.732R_6 + 0.850R_7)X_{acDH} \tag{3.18}$$

$$\frac{dX_{alDH}}{dt} = R_{10} - (0.732R_6 + 0.850R_7)X_{alDH} \tag{3.19}$$

for batch reactors

$$D = 0 \tag{3.20}$$

for fed-batch reactors

$$D = \frac{1}{V}\frac{dV}{dt} \tag{3.21}$$

for chemostat reactors

$$D = \frac{F}{V}, \quad \text{at steady state} \tag{3.22}$$

3.4.4 The model parameters

The model parameters which are to be distinguished from the stoichiometric parameters, are given in table 3.2. The parameters are either taken from literature or estimated directly from experimental work. The estimation procedure is described in Rotbøll and Jørgensen (1992). Most of the parameters are fitted by simulation such that they represent the experimental results. There is no straightforward method to do the parameter tuning in this type of models. The problem is that the model is to be used in both stationary and dynamic conditions. This means that minimisation of a quadratic cost function of the discrepancy between the model and experimental results in stationary conditions might corrupt the representation of the dynamic environment. The approach

Kinetic constant	Value	Kinetic Constant	Value
K_1	$0,465$	K_{i5}	2000
S_1	$0,012$	S_5	$0,36$
K_{e1}	29	S_{e5}	$0,057$
S_{e1}	$0,145$	K_6	$1,13$
K_{i1}	$6,5$	S_6	$0,018$
K_2	$0,26$	K_8	$0,375$
S_2	$0,0005$	K_{e8}	$0,00365$
M_2	$3,5$	S_8	$0,002$
M_{e2}	$3,5$	K_9	$0,029$
K_3	$8,0$	K_{i9}	2700
S_3	$0,08$	S_9	$0,000001$
K_4	$14,0$	S_{e9}	$9,5$
K_{e4}	$1,25$	K_{10}	$0,0023$
K_{i4}	10000	K_{e10}	$0,90$
S_4	$0,000001$	K_{i10}	500
K_5	1400	S_{e10}	170
K_{r5}	$0,3$		

Table 3.2: Model parameters for the model of yeast

is therefore based on trial and error. The parameter estimation issue is further discussed in Lei et al. (1999)

The reason for using the proposed model is the flexibility that it offers. The same model can be used for investigation of batch, fed-batch and continuous operation of the reactor, with the same setting of the parameters. This is the strong point of the model and the ability to represent the switch in the metabolism when saturation occurs in the biomass formation rate R_1, and it can be seen how the flux is directed towards acetaldehyde which will build up and towards production of ethanol. The key assumption is that it is a compound closely related to acetaldehyde which is the signal compound that controls the directions of the fluxes in the cell. This assumption is supported by the experimental work e.g. Postma et al. (1989). Linearization of the process model for continuous operation is performed in Appendix B.1.

3.5 Simulation results

The model described in the previous section is simulated to evaluate the static and dynamic performance with the parameters in table 3.2. The handles which are left for determination are the initial conditions, depending on the type of operation the feed flow rate and the feed concentration of the substrate has to be determined.

The simulation in figure 3.3 is used to show the general trends in the cultivations, in continuous operation the focus is on the accumulation of pyru-

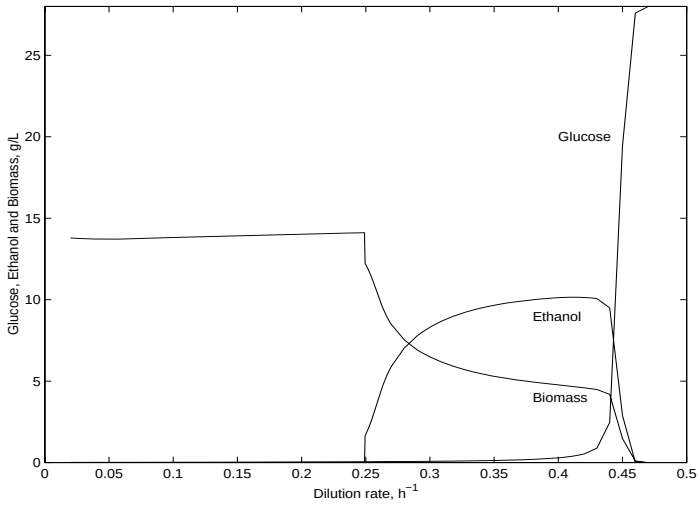

Figure 3.3: Steady states for variation of the dilution rate in chemostat, S_f = 28 g/L, similar to the experiments of Meyenburg.

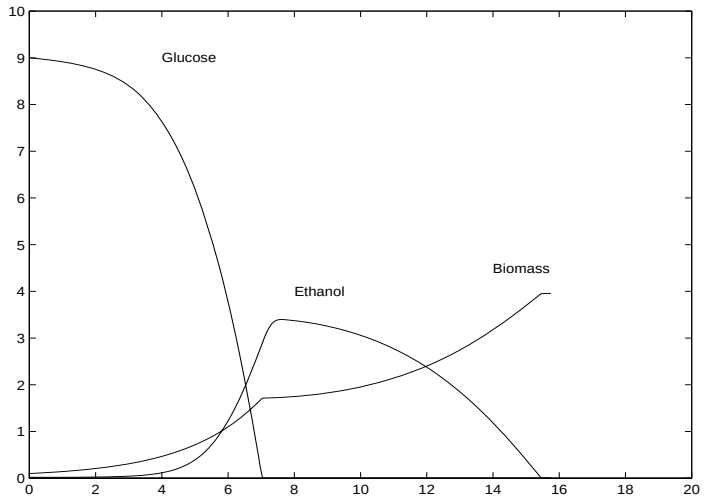

Figure 3.4: Simulation of batch experiment with yeast model, with an initial concentration of the glucose on 9 g/L

vate, acetaldehyde and ethanol. The simulation exhibits the trends reported by Postma et al. (1989), the accumulation of acetaldehyde before the critical dilution rate. A number of simulation verifications have been performed where there is good agreement between the simulation and the experimental results in von Meyenburg (1969), Postma et al. (1989), Andersen et al. (1997). The simulation shown in figure 3.4 of batch experiments show very good agreement with the experimental results, e.g. the lag phase can clearly be seen, but there are some problems with the rate of glucose consumption and the ethanol formation rate in the initial phase of the batch fermentation. These derivations can be explained, in part, with the initial conditions for some of the biomass

components of the model, since they are difficult to determine.

A linearisation of the fermentation process model is shown in appendix B.1

3.6 Conclusion on the model

The described model is an extension of the model of Nielsen and Villadsen (1992) to describe the pathway shift mechanism around the critical dilution rate and thereby to be used in the analysis of both static and dynamic experiments. The flexibility that is introduced by taking this approach is beneficial. By extension of the original model the main features are conserved and new are introduced. This approach could be extended to include growth on more substrates, more detailed description of the glycolysis, R_1, R_2, inclusion of more detailed description of the anabolic reaction e.g. specialized enzymes.

The model shows reasonable agreement with the experimental results and in the following chapters the model will be used for analysis of the operation and for design and analysis of control schemes. It should be pointed out though that the model has some flaws, specially the estimation of the parameters is a problem.

The fact that the parameters are not the best that could be achieved, as mentioned in the section on model parameters, is only a problem if the analysis result is applied directly on a real process.

4

Data analysis of cultivations around critical dilution rate

Steady state operation of yeast cultivations around the critical dilution rate is complicated by the drastic drop in the biomass concentration above the critical dilution rate, therefore control is needed for maintaining operation at the optimal dilution rate. Standard process identification techniques of the input output type will result in bias if applied directly, disregarding the feedback loop. By application of novel identification techniques to yeast cultivations operated in closed loop around the critical dilution rate, it is shown that depending on the data quality improved process identification can be obtained. The experiments were designed to demonstrate the existence of multiple steady states around the critical dilution rate. The analysis of the data shows that there is strong evidence of multiple steady state.

4.1 Introduction

In the biochemical industry the use of continuous cultivation of a microorganism is of significant importance. Yeast is used in many applications for a wide range of products. In many cases the organism is modified through genetic manipulations, with either insertion and/or deletion of genes. Production of these gene products is closely related with the biomass formation. It is known that the optimal productivity is obtained below D_{crit}, but very close to D_{crit}. The critical dilution rate is the dilution rate where the microorganism shifts pathway from oxidative to oxido reductive growth. The shift in the metabolism causes production of Ethanol . At higher dilution rate the yeast cells will produce more ethanol, until wash out of biomass occurs.

At the pathway shift the process shows a very steep drop in the biomass concentration on increases in the dilution rate. In the literature there has been indications that there are multiple steady states around the critical dilution rate Axelsson et al. (1992), of which some will be unstable. Therefore a controller can be used with advantage for obtaining the data at the desired operation point in a set point tracking experiment, To obtain data at the unstable steady state.

Experimental verification of hysteresis phenomenon have been reported in the literature for *Methylomonas* cultivations Dibiasio and Weigand (1981) and

references herein e.g. isothermal auto catalytic reactions and for mold cultivations.

To operate a chemostat around the critical dilution rate with high cell concentration may be difficult in open loop conditions. Even small variations in the dilution rate can cause the pathway switch to be triggered and drastic drop in biomass and production of ethanol is initiated as a consequence, known as overflow metabolism. To be able to obtain experimental data at this operating point it is essential to operate the cultivation in closed loop by measurement of a component for manipulation of the dilution rate. The experiments Møller (1993) where designed to show whether multiple steady state exists in the region around the critical dilution rate. This type of experiments have not been performed earlier for yeast cultivations.

Identification techniques are applied to develop process models based on measured process inputs and outputs. The data are obtained from various experiments, but typically these experiments are run in closed loop to achieve the desired operation point. By applying the standard method for identifying models using closed loop data introduces a bias term since the standard identification techniques does not consider the control loop.

In contrast to the way to model the reactor that was shown in chapter 3, the approach here is to identify process transfer functions based on experimental data.

In this chapter the experimental results obtained by Møller (1993) are applied to the closed loop identification technique by van den Hof and Schrama (1993), For a survey of different approaches for identification for control in closed loop, see van den Hof and Schrama (1995). Their method for improved estimation of closed loop data is presented and input/output models, of ARX types, for continuous cultivation of yeast are estimated. The method is a two step procedure for identification of process models from closed loop experimental data. The data investigated are obtained around the critical dilution rate where operational problems can occur under open loop conditions.

The chapter is structured as follows : first the experiments are introduced with respect to the process and the available data set. Then the closed loop identification method is introduced and applied to the experimental data. Finally the results are discussed and the conclusions is presented.

4.2 Process and Experimental Description

The control problem near the critical dilution rate is complicated due to the shift in the metabolic pathway, to overcome the problems arising from this shift an adaptive controller is used Andersen et al. (1997) when the data is obtained, for details on the operational procedures see Jørgensen et al. (1992). By using an adaptive controller operation at different operating points is facilitated in straight forward manner.

A schematic drawing of the reactor is given in figure 4.1. The controller is using the measurement of reducible gases, Y_{RG}, in the exit gas stream, which

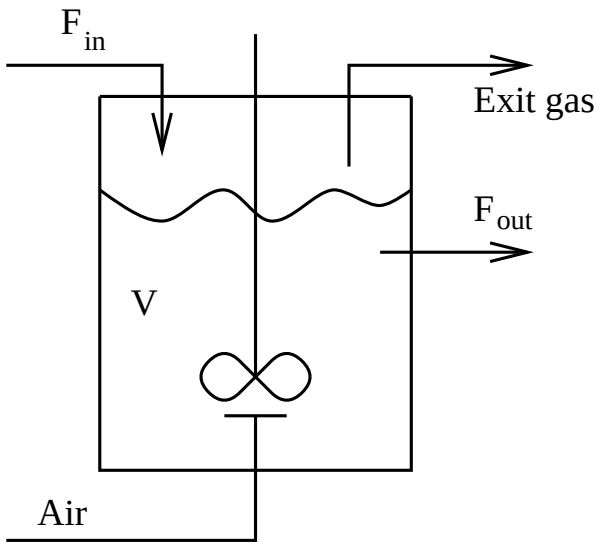

Figure 4.1: Schematic representation of the fermentation process, the feed rate is regarded as process input, the result of the off gas analysis Y_{RG} is the process output

is equivalent to the ethanol concentration in the cultivation broth, for calculating the control input which is the feed flow rate F_{in}. Figure 4.2 show the experimental setup with the control structure.

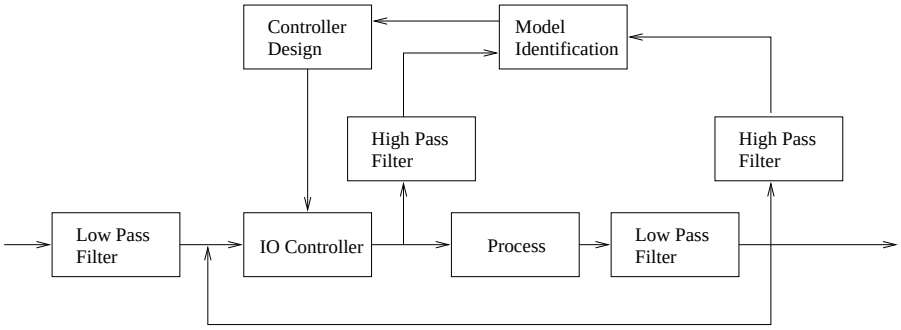

Figure 4.2: Block diagram representation of the closed loop identification setup with adaptive LQ controller

The idea behind process identification is to estimate a process model which relates the process output (Ethanol concentration equivalent) to the process input (feed flow rate). When the process is operated in closed loop, the open loop estimation procedure will result in bias. In the closed loop approach the reference signal for the controlled output is also needed for the identification procedure.

The experimental idea is, by selecting the ethanol measurement as process output and the dilution rate (equivalent to feed flow rate for continuous operation) as process input, to investigate if there exists multiple steady states in the

region around the critical dilution rate. By applying different set points for the measurement Y_{RG} tracking of different operation points is possible. If there exists multiple steady states then they would be captured by this approach since the controlled process input will be adjusted to the new steady state.

4.2.1 The obtained data

The main objective is to improve estimates of the process dynamics around the critical dilution rate, even though the process is operated in closed loop. The use of a two step identification procedure proposed by van den Hof and Schrama (1993) is applied. Ethanol is measured as reducible gas. Measurements are available from four levels of Ethanol 10, 50, 67 and 90 reduced gas units. The sample time is 1 minute. Models will be estimated for all four levels. The levels are shown in figure 4.3. As seen in figure 4.3 the input signal is not

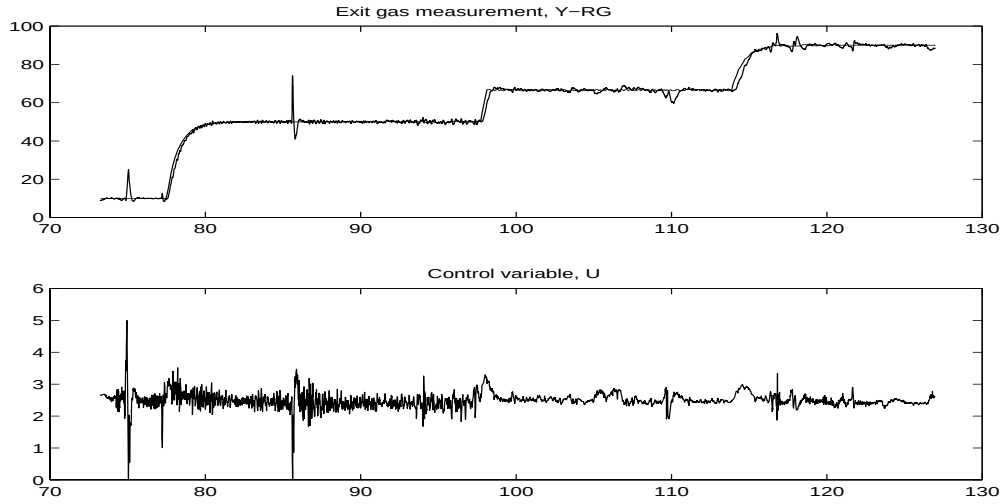

Figure 4.3: Data at the different EtOH levels used in the estimation plotted against time, *h*, the control variable U is the feed flow rate. The operating volume is 8.15 L

shifting level in the same manner as the output. This indicates that the process is operating in a region where there is high sensitivity to small variations in the input level and perhaps even multiple steady states. Another point is that the process input is fluctuating very much for the two lower levels of the Y_{RG} measurement. One reason for this could be that the process is unstable at these operation points and therefore the controller has a much harder work to do at these operation points. Another reason could be that the open loop discrete time process model has a zero on the negative real axis, which is not properly cancelled by the applied control design, at least at $Y_{RG} = 10$. One thing to be noted is the tracking of the set points changes which are performed smoothly and without any problems. Along with the experiment shown in figure 4.3 there is an additional set of data obtained at $Y_{RG} = 10$. This set will be used for

identification at this level since there is no excitation in the reference signal for the data shown in figure 4.3 at this level.

4.3 Estimation Problem

In this section presented the closed loop identification method proposed by van den Hof and Schrama (1993). The main problem in process estimation is to reduce the influence of process noise on the estimated process model. The whole setup of the closed loop is shown in figure 4.4. The experiments

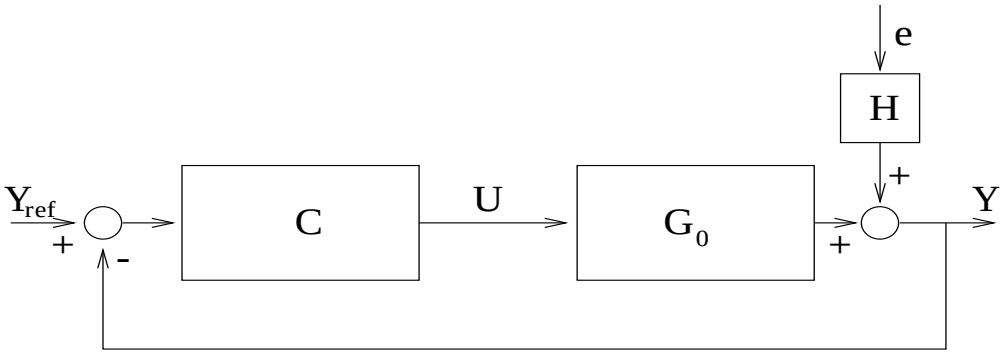

Figure 4.4: Setup of the closed loop control configuration, used in the experimental setting and for understanding the estimation problem

are conducted in closed loop and the general formulation of the estimation problem is based on the transfer function representation of the process, the controller and the noise model. The discrete time process model is described by the following set of equations

$$y(t) = G_0(q)u(t) + H_0(q)e(t) \tag{4.1}$$

where $y(t)$ is the process output (Y_{RG}), $u(t)$ is the input signal (feed flow rate) and $e(t)$ is the measurement noise. $G_0(q)$ are a proper rational function in q, the forward shift operator Friedland (1987). In open loop operation the input signal is perturbed and the process transfer function may be estimated assuming $u(t)$ and $e(t)$ are uncorrelated Ljung (1987).

In closed loop operation, figure 4.4 the control signal $u(t)$ is generated by the controller $C(q)$

$$u(t) = C(q)\left(y_{ref}(t) - y(t)\right) \tag{4.2}$$

where y_{ref} is the reference value for the process output. This signal has to be excited over a frequency range such that the signal can be used in the identification. The problem in this setup from an identification point of view is that the process input and the noise is correlated due to the feedback. This effect makes standard identification inapplicable, due to violation of the standard assumption that the noise and the process input are independent. The prediction

error spectrum for standard open loop identification is given by

$$\Phi_\varepsilon = \left[|G_0 - G(\theta)|^2 \Phi_u + \Phi_e\right] \frac{|L|^2}{|H(\theta)|^2} \tag{4.3}$$

where Φ_u, Φ_e is spectral densities of the input and the noise, respectively, θ is the estimated model parameters. L is some stable filter.

If the open loop technique is used on closed loop data the relevant spectrum is

$$\Phi_\varepsilon = \left[|S_0[G_0 - G(\theta)]|^2 \Phi_r + \frac{|S_0|^2}{|S(\theta)|^2}\Phi_e\right] \frac{|L|^2}{|H(\theta)|^2} \tag{4.4}$$

S_0, $S(\theta)$ is nominal and estimated sensitivity function.

In the following the two step procedure for closed loop identification is presented.

4.4 Closed loop Estimation Method

The closed loop system described by equation 4.1 and 4.2 will be modelled by the following set of models

$$\hat{y}(t) = \hat{G}(q,\theta)u(t) + H(q)e(t) \quad \theta \in \Theta \subset R^d \tag{4.5}$$

where the transfer function $G(q,\theta)$ depends on the real valued parameter vector θ. The nominal process sensitivity function is given by

$$S_0(q) = \frac{1}{1 + G_0(q)C(q)} \tag{4.6}$$

Now using 4.6 and rewriting 4.1 and 4.2 we get

$$u(t) = S_0(q)C(q)y_{ref}(t) - S_0(q)C(q)H_0(q)e(t) \tag{4.7}$$

Since y_{ref} and $e(t)$ are uncorrelated and $u(t)$ and y_{ref} are available from measurements $S_0(q)C(q)$ can be estimate using standard open loop techniques. For the second step the estimated transfer function is used to generate a noise free control signal $u^r(t) = \hat{S_0}\hat{C}y_{ref}$ and inserting into 4.1, which yields

$$y(t) = G_0(q)S_0(q)C(q)y_{ref}(t) + S_0(q)H_0(q)e(t) \tag{4.8}$$

$$= G_0(q)u^r(t) + S_0(q)H_0(q)e(t) \tag{4.9}$$

Then again it is possible to apply an open loop technique for identification of the transfer function $G_0(q)$, since $u^r(t)$ and $e(t)$ are uncorrelated and $y(t)$ is available from measurement. The prediction error spectrum is for the second step given by

$$\Phi_\varepsilon = |[G_0 - G(\theta)]S_0 + G(\theta)[S_0 - S(\beta^*)]|^2 \cdot \Phi_r |L|^2 \tag{4.10}$$

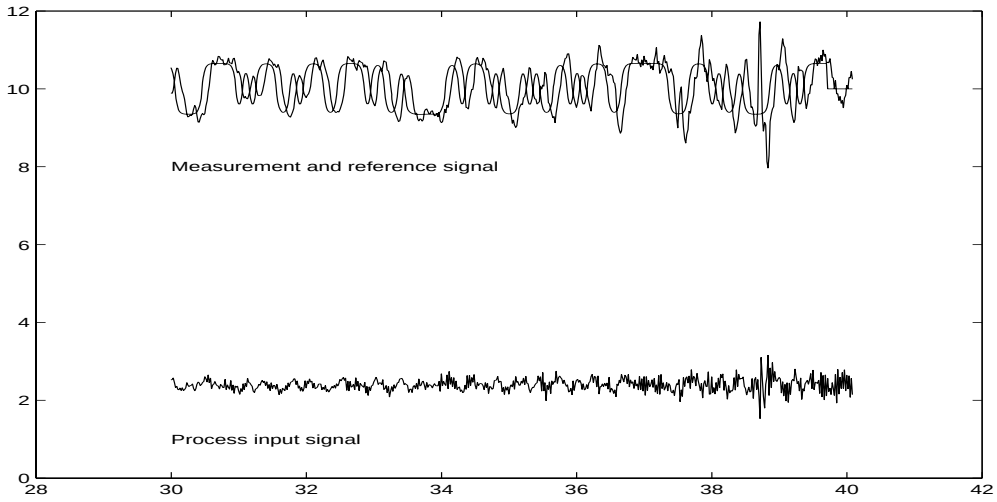

Figure 4.5: Data from experiment at $Y_{RG} = 10$, the time scale is in hours. The feed flow rate is in L/h and the measurement is in relative gas units

Here β^* is the estimated parameters from the first step. If the first step is sufficiently accurate ($S(\beta^*) \rightarrow S_0$) then equation 4.10 reduces to a simple weighted mismatch between G_0 and $G(\theta)$.

One data series is shown in figure 4.5 with the process output (the noisy curve), the reference around 10 (the smooth curve) and the feed rate around $2.5Lh^{-1}$ (the lower curve). A process model is estimated in the two step manner described above,

- Estimate the noise free control the estimated $\hat{u}^r(t) = S_0\widehat{(q)C(q)}y_{ref}(t)$ from y_{ref} and u, through identification of $S_0\widehat{(q)C(q)}$.
- Estimate the process transfer function $\hat{G}(q,\theta)$ from $\hat{u}^r$ and y.

4.5 Results

The result of the estimation procedure is that the process model is now able to provide better predictions than reported in Jørgensen et al. (1992), for some of the operation points. At the levels $Y_{RG} = 10$ and 90 the estimation results were improved compared to estimation only based on the process input and output, whereas at the two middle levels no improvement was obtained, since the first step of the identification procedure failed to produce a reliable result. Figure 4.6 shows that the two step procedure performs better than the open loop. Comparing the first 50 samples it is obvious that the open loop has an offset.

Another interesting aspect is the stability issue. Based on the estimated models it is possible to calculate the zeros and eigenvalues of the process A discrete time process model is stable if the poles are located inside the unit circle. In

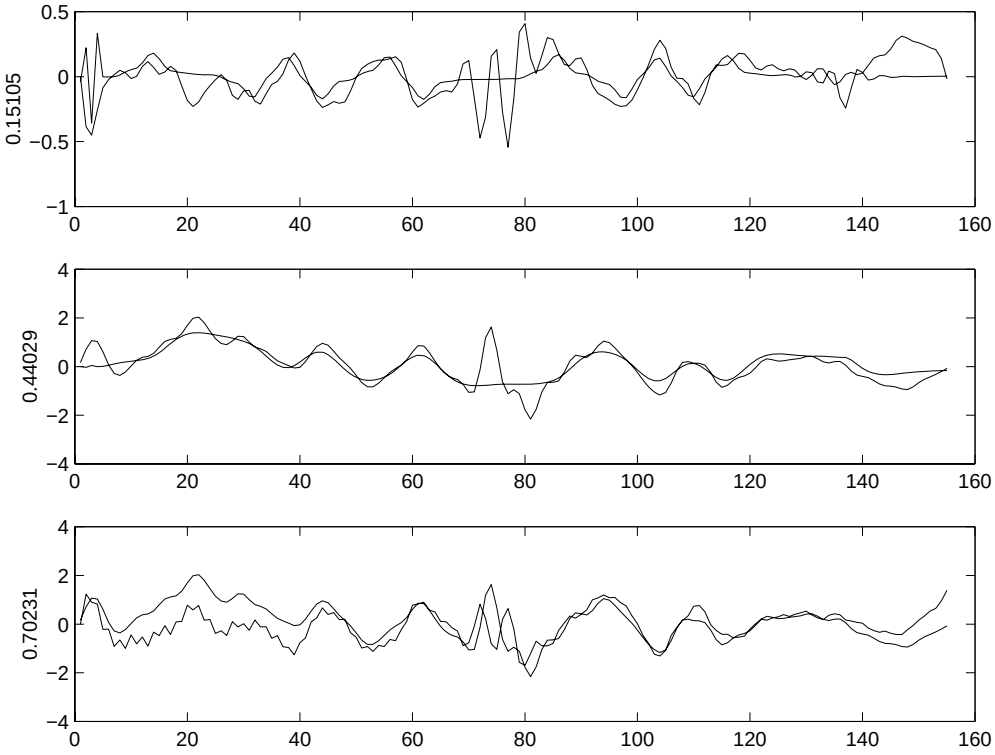

Figure 4.6: Validation of estimation result at level 10. The top figure shows how $u^r(t)$ is fitted to u, based on the data in figure 4.5. The middle figure shows how the second step of the process identification is performing. The lower part is the result generated from open loop identification of the process. The values at the left of the figures is the goodness of fit

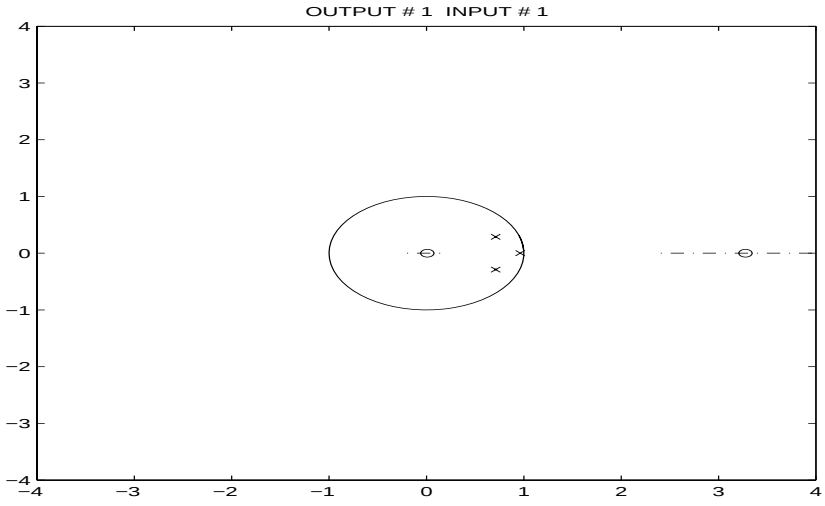

Figure 4.7: Zero and Pole plot,'o':Zero, 'x':Pole, for discrete time transfer function at $Y_{RG} = 10$

figure 4.7 the zeros and poles of the transfer function of the estimated $\hat{G}(q,\theta)$ is plotted. Figure 4.7 shows that the process at the $Y_{RG} = 10$ level is stable, but has an inverse response (a zero outside the unit circle). At the level 90 the results where similar the only major difference was that the zero outside the unit circle was closer to 1. Dynamically this means that the inverse response is slower than at level 10. At the two middle levels the estimation of the noise free control signal is not performing in a satisfactory way. Thus the second step of the estimation fails, compared to the open loop identification.

Open loop identification, based on the available data, shows that the process is stable. This is valid for the frequency range of the perturbated signal, which is limited, see table 4.1 for details.

Instability would be expected to correspond to an eigenvalue with rather slow dynamics, which means lower frequencies than in the available data.

To explain this, one has to look at the assumption of uncorrelated signals in the open loop identification, due to feedback this is not the case as mentioned earlier. The reason for the first step in the estimation procedure to perform badly, can be explained by the limited amount of data available for the estimation and the fact that the quality of these data is rather poor (limited excitation of input signal). The tracking of the set point is illustrated in figure 4.8 which

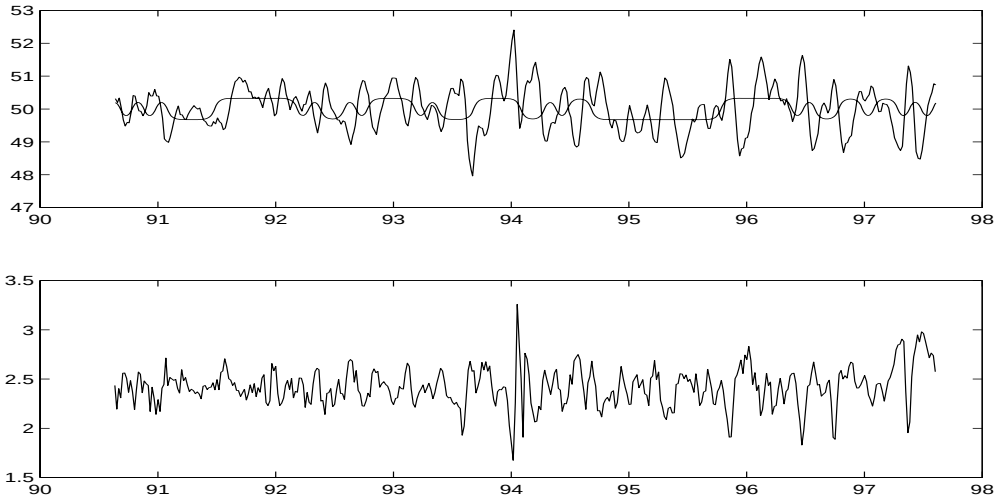

Figure 4.8: Experimental date obtained around EtOH level $Y_{RG} = 50$

show the data from the experiments at $y_{RG} = 50$. It is clearly seen that the process is not able to track the reference signal nearly as well as in figure 4.5. This could indicate that the process in unstable at this intermediate level or the noise level is comparable with the perturbation amplitude. One way to overcome this problem could be to increase the amplitude of the input signal to get hold of the tracking properties of the process.

Level	Eigenvalue OL	Eigenvalue CL	Time constant OL/CL (min)
10	-0.0185 ± 0.0368	-0.0388 ± 0.0019	$54/26$
50	-0.0023 ± 0.8294		435
66	-0.0059 ± 0.0103		169
90	-0.0191 ± 0.0018	-0.0285 ± 0.0054	$52/35$

Table 4.1: The largest eigenvalues of the estimated models at the different Y_{RG} levels. Both open loop and closed loop results. At the level 50 and 66 the estimated models are only based on an open loop estimation

4.5.1 Eigenvalues at the four levels

The estimation of the models at the four different levels was as mentioned in the previous section is not successful in all cases when the two step identification method was applied. To analyse the obtained models the continuous time eigenvalues is calculated, since the eigenvalues are directly related to the time constants of a process.

It is seen from the eigenvalues/time constants that the process is changing its dynamic properties in this region of operation. The difference in the values is around one order of magnitude, which is very large. In figure 4.9 the data from

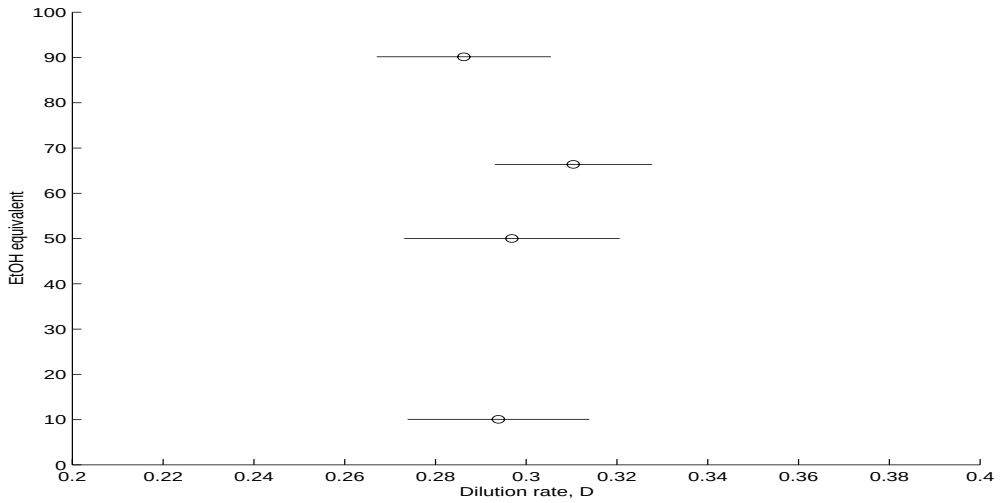

Figure 4.9: Plot of the state space, showing the possibility for multiple steady state, the horizontal lines indicates the standard deviation on the dilution rate, in all cases around $10 - -15\%$.

figure 4.3 is plotted as input output relations.

The occurrence of the multiple steady state in figure 4.9 and the large variations in eigenvalues in table 4.1 is an indication of a fold bifurcation. Since the process at a fold bifurcation point goes from stable to unstable this means that there is an eigenvalue which is zero and moves from the negative half plane to the positive half plane. Since the relation between the eigenvalues and the time

constants is given by equation 4.11.

$$\tau = real(\lambda)^{-1} \tag{4.11}$$

The result of a bifurcation is that the time constant of the process is very large in the region around the bifurcation point. This could explain the large variations in the time constants and their large values around the level $50-66$. It must be noted is that there is a rather large standard deviation of the calculated eigenvalues.

4.6 Conclusions

The application of open loop identification techniques for identification of process models operated in closed loop shows the possibility of getting improved estimated process models by using the two step procedure.

This work has shown that the identification of process models by using closed loop experimental data can be improved if the data is obtained at reasonable experimental conditions. Care should be put to the design of the input reference signal, both frequency and amplitude are of importance. The frequency rang will affect the range of process time constants that can be estimated. The amplitude has to match the noise level, otherwise the result will be corrupted as seen for the two middle levels of Y_{RG}.

The identification problem is in this example complicated by the fact that it is operated very close to the critical dilution rate, where the process exhibits very nonlinear dynamics, and trying to estimate linear models is troublesome, since there will be large variations in the different time constants of the process.

High sensitivity of the measurement Y_{RG} towards changes in dilution rate is seen in figure 4.9, together with the large variation in the slowest estimated time constants is an indication of multiple steady states around the critical dilution rate.

The identification of the yeast cultivation near the critical dilution rate has to be performed for each level, due to the underlying nonlinear structure of the process, where there must be paid attention to the type of input signal applied as reference signal. The frequency range should be broad to capture both the small and the large time constants of the process.

The fact that the controller is adaptive can cause a problem since the estimation procedure assumes that the controller and process is to be considered as constant in a linear sense. This is clearly not the case here particularly at the two levels where the two step procedure fails.

5

Bifurcation analysis Continuous yeast cultivation

The analysis of the developed model is performed by using numerical algorithms to search for the location of the bifurcation points of the continuous cultivation of the microorganism. The analysis reveals the existence of multiple steady states and that the distance between the bifurcation points is increasing with increasing feed concentration. There is a strong need for understanding the underlying mechanism which determine the location of the fold and ways to overcome the operational difficulties which arise at the optimal production point.

5.1 Introduction

In this chapter numerical algorithm for analysing steady state behaviour of differential equations is applied to the continuous yeast cultivation model developed in chapter 3. The idea is to investigate if there exists multiple steady states in the region around the critical dilution rate. It is furthermore possible to investigate the stability of the different steady states.

The motivation behind the analysis of the continuous yeast cultivation is the results obtained in chapter 4 and recent results in the literature.

In many chemical processes there are operational problems due to the existence of hysteresis. This means that multiple stable steady states exist in some region of the operation window. Associated with the stable steady state solution seen as hysteresis curves there exists an unstable steady state located between the two stable states.

The experimental work of Rieger et al. (1983) showed evidence of the phenomenon and the emergence of hysteresis phenomena, which will appear when the dilution rate is varied in one direction and one branch is followed, and when the parameter is changed in the other direction another branch is followed.

Axelsson et al. (1992) showed both by simulation and experiments on Saccharomyces cerevisiae that perturbations can cause the emergence of new steady state conditions, this a unique phenomena of nonlinear processes which is not existing in linear processes.

As seen in the data analysis chapter (chapter 4) the possibility for multiple steady state exists. What is experienced is that the a small change in the bifur-

cation parameter will have a dramatic effect on the out come of the system and simply resetting the parameter will not be sufficient to obtain the original operating conditions. The setting of the experiments will be productostat operation of the reactor. The experiments by Rieger et al. (1983) show that when operation is around the critical dilution rate only small disturbances will cause the appearance of a new steady state. All these indications point in the direction of a fold bifurcation. First the analysis is performed on the simplified model describing biomass growth. Then the results of the analysis of the model presented in chapter 3 is presented and possible explanations for the occurrence of the bifurcation are presented.

Readers not familiar with the terminology and ideas of nonlinear system analysis are recommended to read appendix A as an introduction.

5.2 Optimal Operation vs Fold Bifurcation

Operation of a given process is subject to constraints and is at first optimized with respect to design and secondly the operation is optimized. Given the design and the process constraints the analysis of the model equations will result in a steady state relationship (continuation plot), relating the parameter p with steady state value and its stability. For a large class of processes the result will be a curve with a S-shape where the two outer solutions are stable and the middle one is unstable. The model equations is given by a set of ordinary differential equations

$$\frac{dx}{dt} = F(x,p) \tag{5.1}$$

where x is the n dimensional state vector and p is a process variable. If the analysis is conducted where p is a controllable parameter and optimization of process operation can be formulated as

$$\max_p x_i \cdot p \tag{5.2}$$

such that

$$F(x_0,p_0) = 0$$

In the case of continuous yeast cultivation where the objective is to maximize the biomass productivity. x_i corresponds to biomass and p would be the dilution rate. To determine where the optimal operation point is located, it is seen that the maximum occurs at

$$0 = \frac{dx_i \cdot p}{dp} = p_0\frac{dx_i}{dp} + x_{i0}\frac{dp}{dp} \Rightarrow \frac{dx_i}{dp} = \frac{-x_{i0}}{p_0} \tag{5.3}$$

Meaning that the optimal operation point is located where the slope in the x_i, p-diagram is $\frac{-x_0}{p_0}$ for this type of objective.

The fold bifurcation occurs in the point where the slop in the x, p-diagram is $-\infty$ or ∞. The distance between the optimal operation point and the bifurcation

point is depending on the system at hand but studies of different examples have shown that they are located in the same region of operation. Jørgensen and Jørgensen (1998) show that the optimal operation point can be located in the region where there are multiple steady state solutions. The location of the optimal operating point in the region of multiple steady state solutions can cause operational problems, e.g. disturbances can cause the process to drift towards the other stable steady state, which is undesirable from a productivity point of view.

5.3 Analysis of the operation of continuous cultivations

In recent years there have been some evidence of the occurrence of multiple steady states in the operations of continuous cultivations of microorganisms. In the literature there is some evidence of the hysteresis phenomena as mentioned in the introduction, but the experiments have to be performed in such a way that the phenomenon is revealed. Typically the experiments are performed in such a way that a parameter/variable is changed in one direction over some interval with some predefined steps and with some predefined expectations on the time needed to obtain a new steady state based on a number of residence times typical $5-6$ is considered to be sufficient, which is true if only reactor dynamics is considered. The analysis in this section will show that such experimental conditions will not reveal all the aspects of the system under consideration. Xiu et al. (1998) investigated an unstructured model with three states, biomass, substrate and a product, with inhibition of substrate and product on the growth rate and excess kinetic to account for overflow metabolism. Their analysis show that the model has both multiple steady states and that it can show oscillatory behaviour in some region of the operation regime, which is determined by the substrate feed concentration and the dilution rate. In the region with multiple steady state there are two stable ones and one unstable. The process is operated at the unstable steady state, a perturbation will cause the process to end up at either the one of the stable steady states. The direction of the perturbation will determine which of the steady states will be the result, according to the domains of attraction.

The process is operated close to a fold bifurcation and two arbitrary perturbations (20% and 66% increase of feed concentration) are performed and convergence to new steady state conditions is achieved after $6-8$ resident times. If the perturbation is very small (0.3%) the effect is that the new conditions is reached after something like 40 residence times. This is complicated by the fact that almost nothing is revealed for the first 24 residence times. These transient experiments is not explained in full detail by Xiu et al. (1998). They argue that the time required to obtain steady state is determined by the size of the disturbance, as when the process is operated at the unstable steady state.

The first reason for the long time before the disturbance is seen in the output

is because the disturbance is very small but since the process at this operation point is at a bifurcation point then the time constant of the system is large due to the bifurcation (eigenvalue is zero), which will result in a prolonged response time.

5.3.1 Introduction to bifurcation analysis on Yeast

The primary interest of this chapter is the operation of yeast cultivations. As indicated in chapter 4 and references in the introduction there is experimental evidence that yeast show hysteresis and have multiple steady states around the critical dilution rate, of which one is unstable. To see what can cause the appearance of multiple steady states the following model from Kuhlmann et al. (1998) is considered.

$$\frac{dS}{dt} = -\frac{\mu}{Y}X + (S_f - S)D \tag{5.4}$$

$$\frac{dX}{dt} = \mu X - XD \tag{5.5}$$

The model have two stationary solutions, one $(S,X) = (S_f, 0)$ which is uninteresting, the other solution is found by solving

$$\frac{dX}{dt} = 0 = \mu_0 - D_0 \tag{5.6}$$

$$\frac{dS}{dt} = 0 = -\frac{D_0}{Y}X_0 + (S_{f_0} - S_0)D_0 \Rightarrow X_0 = Y(S_{f_0} - S_0) \tag{5.7}$$

μ is only a function of the substrate concentration and therefore the substrate concentration can be found from equation 5.6 and the biomass concentration is then given by equation 5.7. To be able to determine a bifurcation point the Jacobian of the process equations 5.4-5.5 is needed.

$$J = \begin{bmatrix} -\frac{X}{Y}\frac{d\mu}{dS} - D & -\frac{\mu}{Y} \\ X\frac{d\mu}{dS} & \mu - D \end{bmatrix} \tag{5.8}$$

When the Jacobian is evaluated at steady state conditions the element $J_{2,2} = 0$. Since the condition for a fold bifurcation is that the eigenvalue of the linearized process equations is zero, meaning that there should be either a row or a column of zeros in the Jacobian, meaning that $J_{1,2} = 0$ or $J_{2,1} = 0$. The conditions for having oscillations is that the eigenvalues of the Jacobian is a complex conjugated pair. The conditions in the simple case is that $J_{1,1} = 0$ and $J_{1,2} \neq 0$ and $J_{2,1} \neq 0$. The condition $J_{1,1} = 0$ meaning that since $D > 0$ then $\frac{d\mu}{dS} < 0$ meaning that growth should decrease if the substrate concentration is increased. This will result in $J_{2,1} < 0$ This simple model can not show the oscillations that is referred in the literature as oscillations in glycolysis and oscillations due to synchronization of the cell cycle. The fold bifurcation is possible if the condition $J_{2,1} = 0$ holds, since the growth is on a single substrate the expression

$-\mu = 0$ is only satisfied at $S = 0$ which is an uninteresting point. The condition $J_{2,1} = 0$ corresponds to the growth has a maximum and this can be provided if the growth is inhibited by the substrate. Simple calculation show that the growth kinetic given in Kuhlmann et al. (1998) there will be a bifurcation at the maximum growth rate as predicted by the above analysis. The kinetic is given by

$$\mu = \mu_{max} \frac{S}{K + S + K_i S^2} \tag{5.9}$$

The values of model parameters and coefficients are given in table 5.1

S_f	10	Y	0.5
μ_{max}	1	K	0.03
K_i	0.5		

Table 5.1: Model Parameters and kinetic coefficients, all taken from Kuhlmann et al. (1998)

Steady state solution of the biomass yield

$$X = Y(S_f - S) \tag{5.10}$$

insertion of this in the substrate equation yield

$$-\mu \cdot (S_f - S) + (S_f - S)D = -R + L = 0 \tag{5.11}$$

By plotting the two terms R, L as a function of the substrate concentration, S, it is seen that the steady state solution is given by the intersection of the two curves. The slope of the straight line in figure 5.1 is determined by the dilution rate D.

The two values are $D = [0.6 \quad 0.8032]$, for the lower value of the dilution rate there are three intersections and evaluation of the eigenvalues of the Jacobian reveals that the steady state located at $S \approx 1.3$ is unstable. For $D = 0.8032$ the line L is a tangent to R thus there are two steady states. For larger value of D there will only be one steady state, the one at $S = 10$. Another way to illustrate the location of the fold bifurcation of this simple model is in the bifurcation diagram, see figure 5.2 Stability analysis of the model shows that the middle steady state is unstable, meaning that operation at this steady state in open loop in practice will be impossible.

This type of analysis of simple unstructured models is shown in a number of papers e.g. Dibiasio and Weigand (1981) and Xiu et al. (1998). The value of these results is limited since the model contain little information on cell metabolism.

5.4 Yeast Cultivation

This section will describe nonlinear analysis of the model described in chapter 3. The model describes the growth of *Saccharomyces cerevisiae* also known as

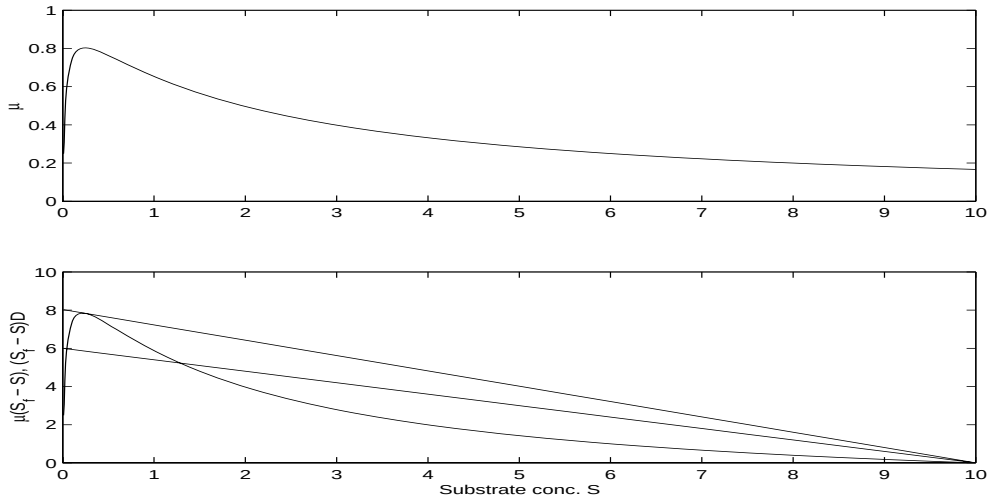

Figure 5.1: Illustration of equations 5.9 and 5.11, with parameters from table 5.1. The upper curve illustrate the growth rate dependence on substrate. The lover figure show the curve R and the two lines represent different values of dilution rate $D = [0.6 \quad 0.8032]$. The intersections represent steady state solutions.

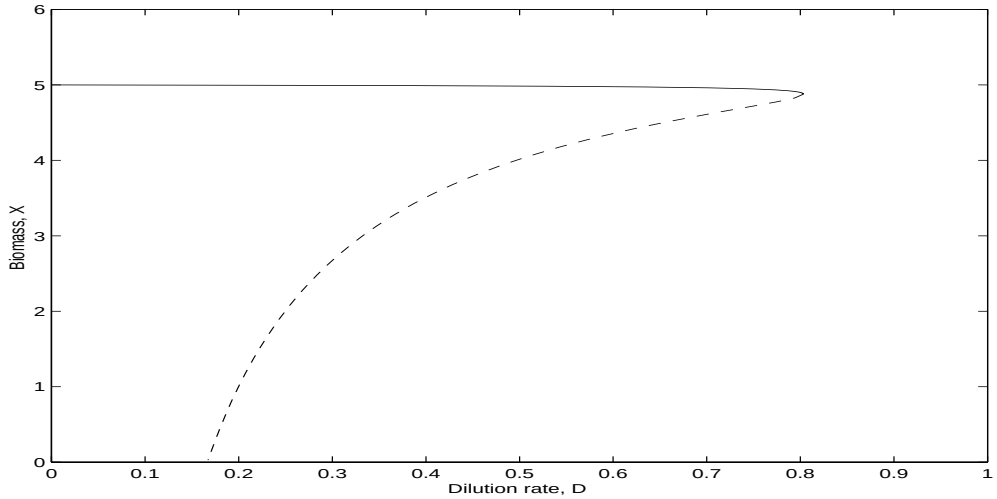

Figure 5.2: Bifurcation diagram for the simple model. The full line denotes the stable steady states and the dashed line the unstable steady states.

bakers yeast. This model exhibits the same type of behaviour as the previous example, but differs largely in complexity. The model will be analysed with respect to variations in dilution rate and substrate feed concentration.

5.4.1 Yeast Model

The model is described with a structured biomass, the structure of the biomass is two enzymes and an active compartment and a structural/passive compartment. The biomass is growing on the main substrate, glucose, which is converted by glycolysis to pyruvate, to acetaldehyde and finally to by product Ethanol. The reaction scheme is as in figure 5.3. The model equations are

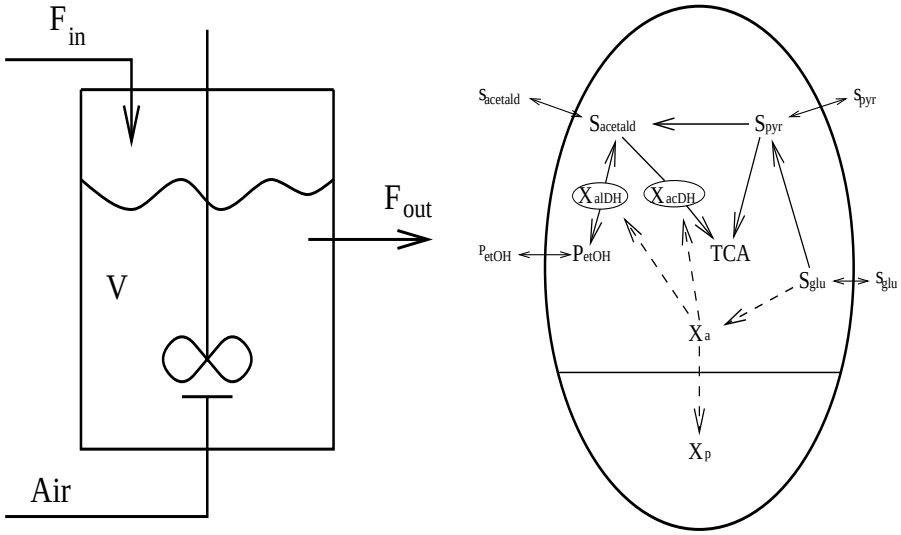

Figure 5.3: Setup of the reactor system and the structure of yeast and the modelled reactions

given in chapter 3.

The question of the level of detail in the modelling of microorganisms is still open, in this model the level of detail was selected with the aim of describing the behaviour around the critical dilution rate. One of the strong points of this model is the ability to be reasonable valid in both fed-batch and continuous operation mode with the same parameter set. In many papers which deals with the modelling of microorganism they use different parameter sets for these two type of operation.

5.4.2 Continuous cultivation of Yeast

The continuous operation mode is used for cultivation at constant conditions and for production of either biomass or some recombinant enzyme. The similarity with the auto catalytic example showed in Appendix A is obvious, biomass is generated by cell division and therefore is auto catalytic. The process con-

trol variables that can be manipulated are the dilution rate and the feed concentration of glucose. The yield of a cultivation of yeast is that approximate 1 g-glucose feed to the reactor can be converted into $1/2$ g-biomass. Another characteristic of a continuous cultivation is that the dilution rate only can be increased up to a certain level then the microorganism shifts metabolism and starts using the glucose for ethanol production, at the cost of biomass production which stabilize at a significantly lower level. If the dilution rate is increased further then the production rate of biomass at some dilution rate is to low compared to the amount of biomass that is removed with the effluent stream. The point where the two rates matches is termed wash out. The interesting point from a biomass production point of view is determined by $(J = D \cdot X)$, the optimal point is located very close to the critical dilution rate.

5.4.3 Nonlinear analysis of continuous east cultivation

The continuation program CONT by Kubicek and Schreiber (1997) is applied to the process model to investigate how variations in the dilution rate will affect the process steady states. The continuation program is constructed in a way that for a given value of the bifurcation parameter the steady state equations are solved, and a step is taken in one direction of the bifurcation parameter. At each step the eigenvalues are calculated and stability changes are checked from step to step. This procedure is applied to the yeast model over the whole range of dilution rates (from zero to wash out). In figure 5.4 the interesting

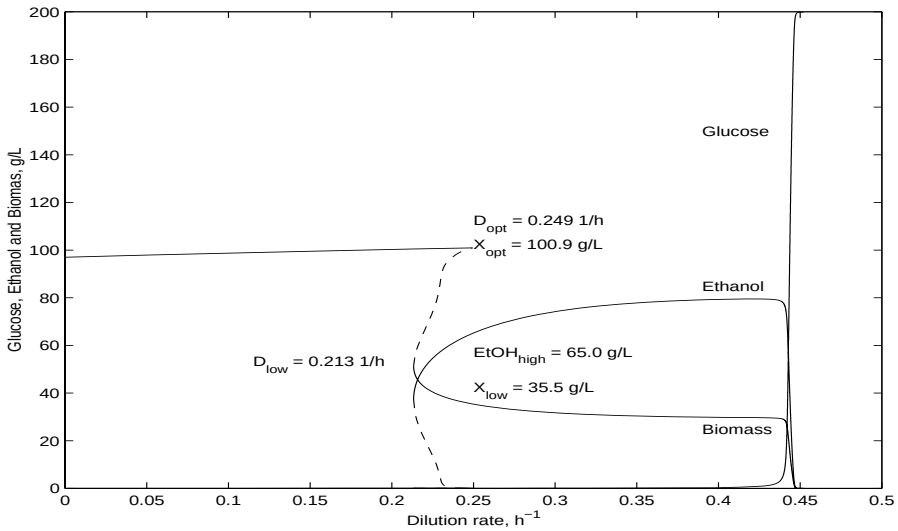

Figure 5.4: Bifurcation diagram with dilution rate as bifurcation parameter, $S_f = 200$ g/L. Full lines represents stable steady states and dashed lines unstable steady states.

fold in the biomass occurs at $D = 0.249461$. The loss in productivity is clear from the biomass curve where the steady state values are 100.9 g/L and 35.5

g/L respectively. This difference gives a loss in production of 65% which is extremely undesirable. And because the eigenvalues at the bifurcation point are close to zero the system is responding slowly (time constants are inverse of eigenvalues). This means that the recovery from a fatal disturbance (e.g. step-up in feed concentration), may take long time. Simulations (not shown) of the model show that the ethanol production is easily triggered by a sudden excess of glucose if the process is operated at the critical dilution rate. If ethanol is present in the reactor the only way to increase the biomass production is to reduce the addition of glucose and let the biomass consume all of the ethanol before the addition of glucose can be increased again. A step-down in feed concentration is not nearly as fatal as step-up because the only consequence of the lack of glucose is a reduction of the biomass level corresponding to the new feed concentration.

The stability issues of the cultivation process is illustrated in figure 5.5 by calculating the eigenvalues of the linearised process equations at each steady state.

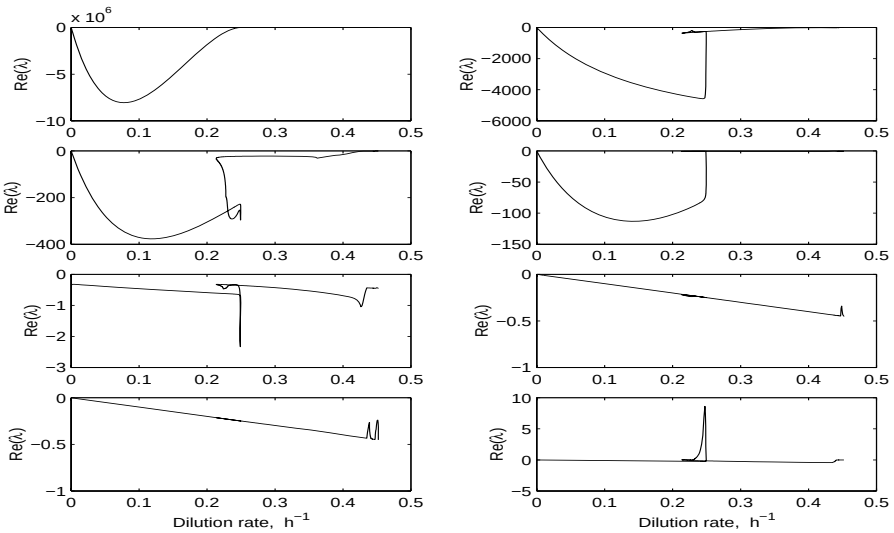

Figure 5.5: The eigenvalues corresponding to the steady states obtained in figure 5.4, only real part of the eigenvalue is shown. The lower right figure show positive real part of the eigenvalues

As can be seen from the calculated eigenvalues there are large variations in the scale of these values. Since the eigenvalues are related to the time constants these will change as the operation conditions change. Not only the stability of the process change around the critical dilution rate, the response and dynamical characteristics is also affected dramatically. A change over three orders of magnitude is seen. The eigenvalue with largest real part is shown in figure 5.6 to illustrate the change around the critical dilution rate.

To investigate the influence of the substrate feed concentration on the stability issues a two parameter bifurcation analysis is performed. In two parameter

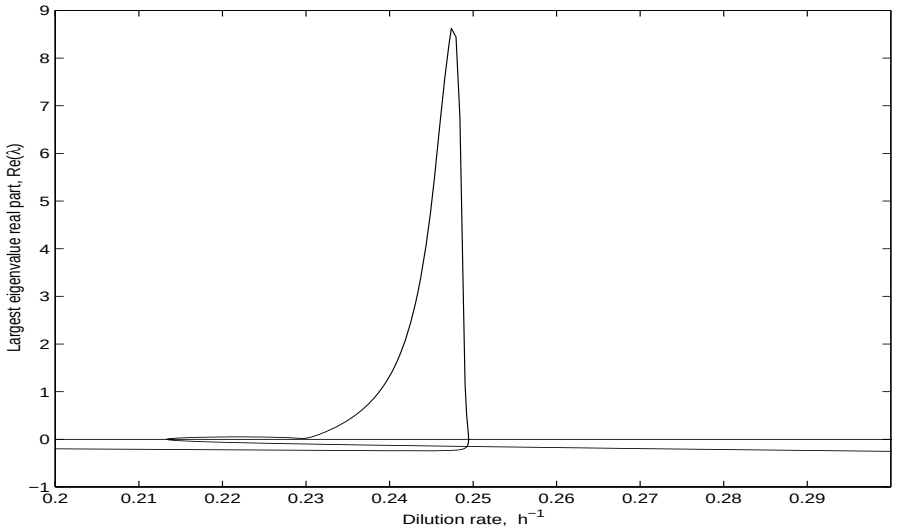

Figure 5.6: The largest eigenvalues corresponding to the steady states obtained in figure 5.4 around critical dilution rate, only real part of the eigenvalue is shown

bifurcation analysis the algorithm tracks the fold bifurcation points as the parameters are varied in the parameter space. As parameters for this analysis is dilution rate and substrate feed concentration. The tracking of the fold points from figure 5.4 is illustrated in figure 5.7

Area one deserves a comment. The upper part of the area corresponds to the production regime where ethanol is produced, whereas the lower part corresponds to the case where only biomass is produced. Figure 5.7 show that the risk for experiencing operational problems increases as the substrate feed concentration is increased since the gap between the bifurcation points measured along dilution rate increases.

In area three there is a double S-curve in the one parameter continuation resulting in five steady state solutions.

The time evolution of a disturbance around the fold bifurcation are shown in figure 5.8. The simulation of the chemostat is performed at two different dilution rates $D = [0.249275 \quad 0.2493]$, the initial conditions are the same in the who simulations. The result of this small difference in the dilution rate is first seen after $75h$ and steady state is not yet reached at $150h$. To explain this from figure 5.4, one has to realize that both dilution rates are below the fold point. The explanation is that the difference in the dilution rate causes the state vector of the process to enter the domain of attraction of the low producing stable steady state. To explain the result on basis of the biochemistry and the kinetic expressions of the model, the rate expressions is calculated and scaled with the active part of the biomass, in order to get the pure kinetic expressions with the inhibitions. The result of these calculations are shown in figure 5.9. The rate expression shows that the reaction from acetaldehyde to CO_2 reaches saturation and the overflow towards ethanol is triggered, resulting in an increased

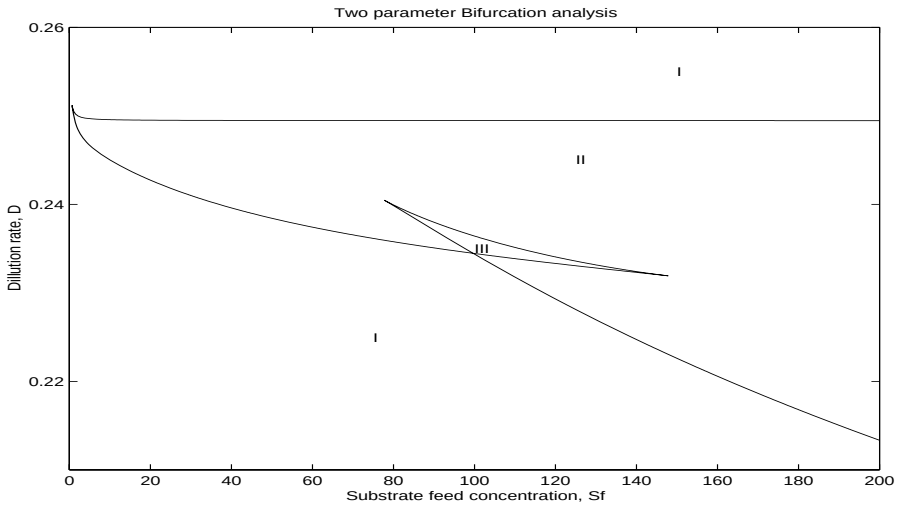

Figure 5.7: Tracking of fold bifurcation points. The areas in the figure corresponds to: I : One single stable steady state. II : Two stable steady states and one unstable steady state. III : Three stable steady states and two unstable steady states.

flux through the pathway and accumulation of acetaldehyde. From figure 5.8 it can be seen that glucose is building up very slowly in a very low amount. This will cause an increase in the flux through the system towards acetaldehyde, until the reaction R_4 (acetaldehyde to CO_2) is saturated and accumulation of acetaldehyde is taking place. The accumulated acetaldehyde will then activate an increased uptake of glucose which then will be converted to ethanol, since this reaction is also activated by acetaldehyde.

5.5 Conclusion

The model for continuous cultivation of yeast developed previously has been shown to have multiple steady states. This is in good agreement with experimental results such as Axelsson et al. (1992), Rieger et al. (1983) and Dibiasio and Weigand (1981) and the analysis in chapter 4.

This is the first time a structured model have been analysed to show multiple steady states, all previously reported studies have used simple unstructured models, such as section 5.3.1.

The identification of the multiple steady states in the model is performed by using numerical techniques and the two parameter investigation of the model showed that the distance between the two bifurcation parameters is increasing for increasing substrate concentrations. This means that the risk of operating around the upper bifurcation point is increasing, due to the large gap between the high producing stable steady state and low producing stable steady state, 65% loss in productivity have been shown.

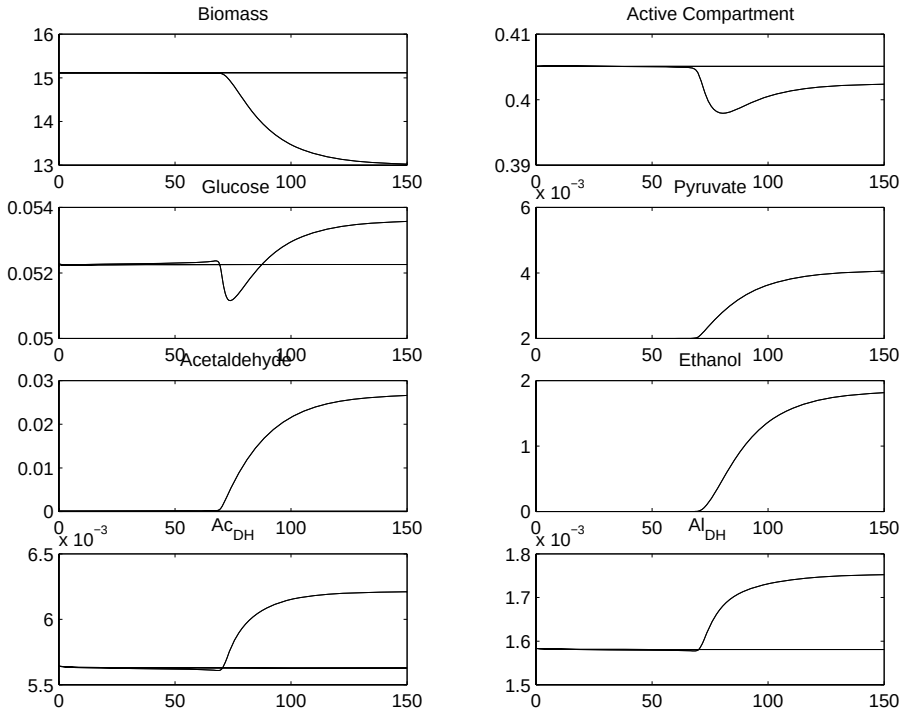

Figure 5.8: Simulation of the continuous cultivation of yeast at different dilution rates, around the fold bifurcation, $D = [0.249275 \quad 0.2493]$. The feed concentration $S_f = 30g/L$

The area III in figure 5.7 is in my opinion rather speculative and might be an artifact of the model. Studies have shown that if some of the model parameters is changed then area III disappears.

The simulation of the model perform as expected both from knowledge of process behaviour around bifurcation points and similar simulations performed by Xiu et al. (1998), that is if the initial point (here initial point covers both state vector and bifurcation parameter) is located close to a bifurcation point then the dynamics of the system is very slow and the system is very sensitive to small perturbations and their direction.

To understand how the fold bifurcations arises one has to look at the determinant of the full Jacobian which for 8 states have 40320 elements, but due to zeros in the Jacobian the determinant reduces to approximately 700 elements, which is impossible to analyse to find an explanation for the fold.

In the simple case the fold occurs as a coupling between the reactor and kinetics.

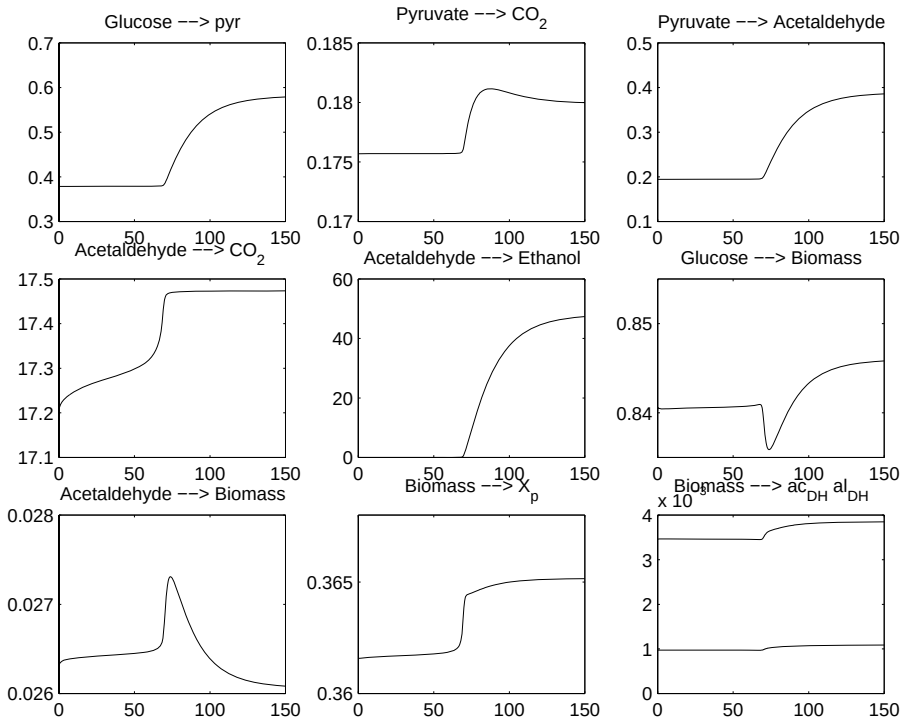

Figure 5.9: Calculated reaction rate from simulation of yeast at the dilution rates, D=0.249275, around the fold bifurcation. The rates expressions are taken from chapter 3 in equation 3.1-3.10, without the influence of the biomass components, X_a, X_{acDH}, X_{alDH}.

6

Control Theory and Application

In this chapter control strategies for the yeast cultivation process is developed and investigated. The different modes of operation impose different control problems and therefore different control strategies are investigated. One of the characteristics of the fermentation process is that there is only a limited number of control handles available and most of these are not included in the analysis and design of the control strategies, since they are assumed to be controlled individually and to such an extent that the controlled variables have no effect on the growth of the biomass. The controllers investigated here are a number of nonlinear control techniques and a new scheme are introduced and discussed

6.1 Introduction

The biotechnical process industry is a rather novel industry compared to the chemical industry and the competition is less pronounced. The lack of the competition due to patents and the tight regulations pharmaceutical products result in what could be termed as *operation as yesterday*. Why change the operation when the process is making profit? It should be kept in mind that there are tight regulations imposed by health authority e.g. national, European and the American counterpart Federal Drug Administration.

What is the major difference of these two industries ? Well there are many, but from a control point of view, the time constants are different, fermentation process are typically rather slow. Another issue is the lack of on-line measurements for relevant process states.

Typically measurements as CO_2, O_2 and EtOH can be measured on-line and sometimes biomass as dry weight measure. The number of available control handles is also small since the physio-chemical measurements, e.g. reactor temperature, pH, DOT etc., is assumed to be controlled by cooling jacket temperature, acid/base addition impeller speed. Which leaves only a few remaining possibilities, being dilution rate and feed concentration. Here dilution (feed) rate is used for control and feed concentration is seen as a disturbance in the test simulation.

Much effort is put into monitoring and fault detection of the processes based on available process measurements and development of strategies to handle identified failures. This chapter describes how to define, analyze and solve the

control problem for the different modes of operation. First the time varying type of operations (batch and fed-batch) is treated in section 6.2 and 6.3. Then the continuous operation is handled. Emphasis is on the control of fold bifurcation arising in continuous operation, which is an unsolved control problem. In section 6.6 control of fold bifurcations is discussed and a novel control law is proposed. This control is applied to the continuous yeast cultivation, evaluation on performance is simulated with disturbances in section 6.7.

6.2 Batch Operation

Operation in batch mode is an important part of the whole operation of a total cultivation since it is the initial phase in the possible long run of the total cultivation period. The purpose of the batch phase is typically to generate biomass for further cultivation in either fed-batch or continuous operation. There are many aspects that will influence how the batch cultivation will behave. The characteristics of batch operation are that no substrate is being added to the cultivation broth, but small amounts of different substances might be added, e.g. for pH control, anti foam. The main handles to influence the operation of a batch is reduced to temperature, pH and initial composition of the broth. It is assumed that the temperature and pH are well controlled and will have no effect on the remaining process variables. Depending on the purpose of the cultivation it is possible of formulate an optimization problem and solve it by using the technique described in section 6.3. The characteristics of batch processes are that it is time varying since the microorganism is consuming the available substrate. When all the energy sources are consumed or some other limitation is meet, the process behavior will change. In the yeast case the model described in chapter 3 is simulated in batch conditions and the result were shown in figure 3.4. In table 6.1 the (not) using of batch is listed.

Since the batch and fed-batch case have similarities with respect to being time varying, the discussion and the concept for solving the control problem is similar. Typically the free variable in batch cultivation can be the initial conditions and the time to add an inducer to initiate production of the desired product and maybe the temperature profile if it has an effect. In mathematical terms this could be formulated as 6.1

$$\max_{Temp(t),T_{add},X_0} \Phi(x(T_f), Temp(t)) \tag{6.1}$$

such that

$$\frac{dx}{dt} = F(x, Temp(t), T_{add}), \quad X_0 = x(t=0) \tag{6.2}$$

Possible with constrains on both inputs, state variables and initial conditions.

Advantage	Disadvantage
Versatile: can be used for different reactions every day.	**High labor cost:** skilled labor is required.
Safe: Can be properly sterilized. Little risk of contamination and strain mutation	**Much idle time:** sterilization, growth of inoculum, Cleaning after the cultivation.
Complete conversion of substrate is possible.	**Safety problems:** when filling, emptying and cleaning.

Table 6.1: Advantages and disadvantage of batch operation, from Nielsen and Villadsen (1994).

6.3 Fed-batch operation

Fed-batch operation is characterized by the addition of substrate to the tank reactor, but no withdrawal from the broth. The fed-batch process is used for fermentation in the bio-chemical industry, for production when the substrate is inhibiting the desired product, since the substrate concentration can be kept low compared to the feed concentration. The dynamics of the process causes the control problem to be time varying by nature, due to the increasing volume. The problem with this type of operation compared to continuous operation is that there is no steady state, due to the varying volume. The problem of controlling states as the temperature or pH, typically having fixed set points is assumed to be solved to a degree such that they will not affect the metabolism of the microorganism. The challenge of operating a fed-batch cultivation is to determine the feed profile, due to the complexity of the coupled chemical reactions in the microorganism, this is not a simple task. Typically the feed profile has been calculated off-line, and then applied to the fermentation. This approach is far from optimal, due to the fact that model uncertainty and process variations may cause break down of the process, e.g. by overfeeding the reactor, which will result in the overflow metabolism producing ethanol.

One way to overcome this problem is to apply closed loop control, such that the process is controlled at a certain state. The problem in this is that most states are varying in time. To solve this one can consider to control the process along a prespecified trajectory. This imposes a problem: What if the process is far from the specified trajectory ? To answer this question process knowledge is needed, since the answer is depending on the actual location of the operation point and performance criterion. Control around a trajectory is preferable to open loop operation, since more consistent production can be obtained from batch to batch.

One approach to deal with these problems will be model predictive control (MPC) on which an optimization algorithm is imposed. This problem can be solved by different approaches, Biegler is solving the differential equations and performing the optimization simultaneously Biegler (1984), for differential algebraic systems Cuthrell and Biegler (1987). The basic idea in his work

is to use orthogonal collocation to discretise the differential equations in time on finite elements and in the optimization to calculate the process input, such that an objective function is maximized or minimized. The advantage of this approach is the computation time, which is smaller than for strategies not using orthogonal collocation, but simply use simulation strategies embedded in an optimization algorithm Biegler (1984). Another approach to calculate the optimal feed profile is to use iterative dynamic programming Luus (1993) and Luus (1994). Both approaches are not directly suitable for application in an experimental environment, due to the high demand on calculation power.

In the following different approaches to solving the control problem for fed-batch operation is discussed. One appealing approach is Model predictive control but the method have also a number of drawbacks, this is discussed in section 6.3.1. In this section the dynamic optimization method by Biegler (1984) is presented.

The issue of robust operation of fed-batch cultivations is discussed in a recent paper Kuhlmann et al. (1998), the approach is evaluated in section 6.3.3. By robust operation is meant that control accounts for model parameters are uncertainty. The consequence is illustrated by the application of nominal optimal trajectory to the process with the worst case parameters set which result in a loss in productivity of about $2/3$ in open loop.

6.3.1 Model predictive control

Model predictive control is a conceptually simple control scheme for solving complex control problems. The reason for the increasing interest is the development of faster computers, which are necessary due to the relative extensive calculations. Another advantage is the simple way of implementing constrains, which can be difficult to take into account in other control design schemes. The principle behind model predictive control is, that the control action is obtained at time t and state x by determining on-line the (open-loop) control $\hat{u}$ by solving a (finite time) optimal control problem, over the interval $[t, t+T]$ and setting the control to $\hat{u}$. Repeating this calculation at every sampling interval yields a feedback control, since $\hat{u}$ is determined from the current state $x(t)$. The optimization is performed such that temperature and process dynamics are handled by the algorithm. There is, due to the finite time, no guarantee that this scheme is stable unless that there is specified some constraint for that purpose the optimization. The stability issue is discussed in Rawlings and Muske (1993), and for an introduction to model predictive control Garcia et al. (1989) discuss many issues of the theory e.g. , advantages and limitations. Generally the largest problem in NLP is getting the global optimum and to guarantee convergence of the optimization

Stability of receding horizon control for nonlinear systems is proven by Mayne and Michalska (1990), robustness of constrained systems Michalska and Mayne (1993) and several of the contributions in Clarke (1994) covers many different aspects concerning the theoretical problems of MPC.

Why is MPC not implemented on every system today ? There are many answers to this, first the transfer of the technology to the vendors and to the process industry, this takes time, the are some application at refineries. Secondly the issue of reliable prediction models for the bio technical industry is generally not available. The MPC method is based on full state information and this is presently not the case for cultivation processes. Finally the people who is developing the cultivation processes are generally reluctant to implement this (or any) type of control, the reason for this is partly due to control people (lack of convincing arguments) and partly the people related to process development (lack of control knowledge).

6.3.2 Dynamic optimization - problem and solution method

The optimization of the fed-batch process is important due to the high capital cost of process equipment and the expensive substrate in order to optimally utilization of allocated resources. In general the formulation of the optimization problem can be broken down into several steps:

- Formulation of an appropriate criterion (objective function) to assess the performance of the reactor. Typically this may be an economic performance criterion, designed to balance the conflicting objectives of maximum productivity, yield and minimum time and utility cost.
- Selection of control variables; initial concentrations, feed addition, sample time, ect.
- Specification of constrained and bounds on control variables and states, due to equipment limitations and other external conditions.

In general the fed-batch process is described by a set of differential equations

$$\frac{dx}{dt} = F(x,u,t) \tag{6.3}$$

with some initial conditions $x(t=0)=0$, where $F(x,u,t)$ typically is a nonlinear function, As an example see chapter 3 for a description of a biological process model. The problem in dynamic optimization is; given an objective function, $\Phi(x,u,t)$ and a process model 6.3, minimize (or maximize) the value of the objective by altering the process input u according to process and the constrained on the states and the inputs. The Mathematical formulation

$$\min_u \Phi(x,u,t) \tag{6.4}$$

such that

$$\frac{dx}{dt} = F(x,u,t)$$
$$x_l \leq x \leq x_u$$
$$u_l \leq u \leq u_u$$

In addition to these constrained there can be additionally soft and/or hard constraining equations, typically given by

$$g[x,u,t] \geq 0$$

$$h[x,u,t] = 0$$

The problem of applying a standard optimization algorithm is the differential equations. These equations have to be evaluated at each sampling instant, because the input only can be changed at the sample instant, the differential equations has to be solved at each step in the iterations of the optimization problem. One way to deal with this problem is to discretise the differential equations in time. By doing this, constrained will only be algebraic equations, which can be handled by an optimization algorithm, such as SQP. In this approach the discretization will be done by orthogonal collocation on finite elements, the approach is explained in detail in section 6.3.2.2

6.3.2.1 The Objective function

The objective function Φ is an important function for the formulation of the optimization problem and therefore important for the result of the optimization. As mentioned the objective function will typically be based on some economic criterion, which are designed such that the different conflicting objectives is included. The following simple objective function will give the maximal productivity in the shortest time possible;

$$\Phi_1 = \frac{P(T_f)V(T_f)}{T_f + T_d} \tag{6.5}$$

$P(T_f)$ and $V(T_f)$ represents product and volume at the end of the operation period, T_f, T_d is the down time of the equipment due to cleaning and maintenance, which is important for short term cultivations but can be eliminated for long term cultivation. A limitation is that this objective (ϕ_1) does not include the consumption of substrate. One way to deal with this problem is to include the substrate consumption into the objective function.

$$\Phi_2 = C_1 P(T_f)V(T_f) - C_2 T_f - \int_0^{T_f} (C_3 s_1 + C_4 s_2) F\, dt \tag{6.6}$$

The terms in equation 6.6 describes the value of the product, cost of process equipment including general utility costs and the cost of the added substrates The cost of substrate might be neglected if the value of the product is high in comparison. If this is the case then the objective in equation 6.5 is preferable. By applying this one will also avoid the instant of selecting the right cost parameters in equation 6.6, C_1 and C_2 thereby the solution of the problem will be general and not depend on different values of cost parameters.

6.3.2.2 Discretization of differential equation

The main idea in orthogonal collocation is to approximate the differential equations with an orthogonal polynomial, such that the error between the approximation and the real value is minimal. The main reference in the field of orthogonal collocation is Villadsen and Michelsen (1978). In this approach the time discretization is performed on finite elements corresponding to the length of the sample time. This approach is similar to that of Biegler, except that the location of the knots in this method is given initially, where Biegler in some papers uses the location of the finite element knots as an optimization parameter. The solution to the model given by equation 6.7, will be approximated by an N'th order polynomial, which will be given as equation 6.8

$$\frac{dx}{dt} = F(x,u) \tag{6.7}$$

$$X_N(t) = \sum_{i=0}^{N} a_i \, l_i(t); \tag{6.8}$$

where $l_i(t)$ is a Lagrange polynomial defined as in 6.9

$$l_j(t) = \prod_{j=0,i}^{N} \frac{t - t_j}{t_i - t_j} \tag{6.9}$$

With this approximation it is possible to calculate the residue at the collocation points. The location of the collocation points are chosen as the zeros of an N'th order Legendre polynomial defined from 0 to T_{sample}. By substituting equation 6.8 and 6.9 into 6.7, the residual can be written as

$$R(t_i) = \left.\frac{dX_N}{dt}\right|_{t=t_i} - F(X_N(t_i), u(t_i)) = 0 \tag{6.10}$$

The reason for preferring the finite element collocation over global collocation is due to the sampling time which is constant, and the global collocation may be inefficient due to the large variations of the states, that will occur in the fed-batch process. The principle in finite element collocation is shown in figure 6.1

6.3.2.3 Selection of collocation points

Typically the collocation points will be selected as the roots of a Jacobi polynomial, $P_N(\alpha,\beta)(t)$, which fulfills the orthogonality property,

$$\int_0^1 t^m (1-t)^{\alpha} t^{\beta} P_N^{(\alpha,\beta)}(t) = 0, m < N \tag{6.11}$$

Tieu et al. (1995) discuss different collocation methods, and conclude that end-point collocation is preferable, in comparison with the interior point methods presented in Biegler (1984). For end-point collocation it can be shown that the optimal Jacobi polynomial is $P_N^{(1,0)}(t)$.

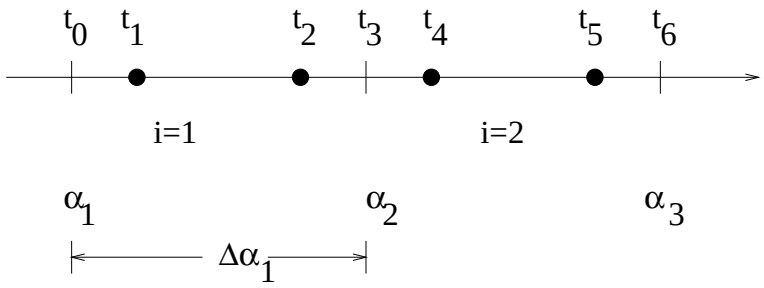

Figure 6.1: Finite element collocation, discretization points for the state profile. The time axis is split into a number of finite elements, these elements is then discretised by collocation determined by the Jacobi polynomial

With the application of end-point collocation the optimization problem can be rewritten into the following nonlinear programming problem, equation 6.12, which can be solved by an NLP algorithm such as SQP.

$$\min_{a_i,u} \Phi(X_N, u, t) \tag{6.12}$$

such that

$$R(t_i) = \left.\frac{dX_N}{dt}\right|_{t=t_i} - F(X_N(t_i), u(t_i)) = 0$$

$$x_l \leq x \leq x_u$$

$$u_l \leq u \leq u_u$$

$$g[x,u,t,p] \geq 0$$

$$h[x,u,t,p] = 0$$

As decision variable is the number of collocation points within each finite element, N. It should be determined such that the results are sufficiently accurate and the computation time is not too large.

The time discretization increases the size of the optimization problem due to the residual equations, one the other hand are the differential equations and the optimization problem solved simultaneously .

6.3.2.4 The residual equation

Given the model 6.3 and the collocation procedure in section 6.3.2.2, the residual equations will be given. The residuals are evaluated to zero at the collocation points. The right hand side of the model is for simplicity of the notation just referred to as $f_i(x,u)$. In the model, equation 6.3, there are N_{eq} differential

equations, resulting in N_{eq} residual equations

$$R_1(t_j) = \sum_{i=1}^{N+1} a_{i,1}\dot{l}_{i,1}(t_j) - f_1(x_N(t_j),u) = 0 \tag{6.13}$$

$$\vdots = \vdots \tag{6.14}$$

$$R_{N_{eq}}(t_j) = \sum_{i=1}^{N+1} a_{i,N_{eq}}\dot{l}_{i,N_{eq}}(t_j) - f_{N_{eq}}(x_N(t_j),u) = 0 \tag{6.15}$$

at each of the $N+1$ collocation points. Thus results are $N_{eq}(N+1)$ equations, which are solved to determine the coefficients $a_{i,k}$. Because the functions $f_i(x,u)$ are nonlinear the parameters $a_{i,k}$ has to be found by an iterative procedure, e.g. a Newton search.

The method have been applied to a fed-batch cultivation of a fungi, the result of this is explained in chapter 7.

6.3.3 Robust operation of fed-batch

Kuhlmann et al. (1998) applied dynamic optimization to a nominal fed-batch model and applied the feed trajectory one a worst case situation and the result was disappointing, productivity was reduced by $2/3$. The worst case situation is a change of the maximal growth rate μ_{max} and the kinetic constant K, both by 10%. The model they used for the investigation is a simple unstructured model with substrate inhibition on growth rate. Optimization in the nominal case will result in maximal growth rate is obtained when this high growth rate is not attainable in the worst case situation substrate will accumulate as a consequence. By taking the variation in the process parameters into account in the optimization result in a trajectory optimizing the worst case situation. This trajectory result in bad performance of the nominal case to overcome this problem a feedback control loop is introduced to allow for online adjustment of the feed rate. Simple proportional feedback is used and result is very promising for both nominal and worst case situation. The conclusion is that the control around a specified trajectory is desirable compared to open loop operation and attention is put to the uncertainty in the optimization.

Another approach to control fed-batch operation is the work by Akesson et al. (1997), where there is applied pulses (both up and down) to the feed rate and then the oxygen response is analyzed to give indication on the substrate requirement of the microorganism. The analysis is based on knowledge of the respiration of the organism. Based on this analysis the feed rate is adjusted by simple rules. The approach have shown to be applicable for fed-batch cultivation, with a number of different situations.

6.4 Continuous cultivation

The operation in continuous mode for fermentation is done for long term production with genetically stable organisms. The problem with long term production with microorganisms is that with a certain probability the cells will mutate and the production capacity is lost. Another problem with continuous cultivation with similar effect as mutations is contamination. If there is two different microorganisms in a reactor they will compete for the available energy sources and the one with the highest growth rate will out growth the other and eventually dominate the reactor. For characteristics of the continuous operation of cultivations processes see table 6.2.

The general characteristics for the continuous operation of a fermentor is that the volumetric flow in and out of the tank is balanced such that the total volume is constant. The result is that it is possible to obtain a steady state if the feed flow rate and composition to the fermentor is kept constant.

If the process is operated close to the critical dilution rate, the process will have multiple steady state, as shown in chapter 5, and a number of disturbances can cause the production to be lost, as shown by Axelsson et al. (1992).

Advantage	Disadvantage
Work all the time: High utilization of reactor, short down time compared to operation time. Low labor cost. **Often Efficient:** Due to the auto catalytic nature of the microbial reactions the productivity can be high. Automation may be very appealing. Constant product quality	**Often disappointing:** promised continuous production for month fails due to a) infection due to improper sterilization of feed streams b) spontaneous mutation of microorganism to non producing strain (e.g. strain instability). **Very inflexible:** Can rarely be used for other productions without substantial retrofitting. **Down stream:** the whole down stream process equipment must be designed for low volumetric rate, continuous operation.

Table 6.2: Advantages and disadvantage of continuous operation, taken from Nielsen and Villadsen (1994).

To define the control for the continuous case, a specification of the operation point is needed. Here the scope of cultivation is defined as obtaining maximal productivity of the biomass, determined by $D \cdot X_{biomass}$. From the bifurcation diagram in figure 5.4 it can be concluded that the optimal point is located very close to the fold point. As mentioned in the introduction the dilution rate is used as an actuator and the CER is used as measurement of the process state.

Closed loop operation is needed to reject disturbances in feed concentration and reduce the influence of noise.

6.5 Controller design

The operation of continuous yeast cultivations is complicated by the fold bifurcation which was shown in chapter 5. The optimal operation point is located in the region where there exists multiple steady states.

There are basically two ways to deal with this problem. One is to back off from the optimal operation point and accept the loss in productivity. The back off has to be sufficiently large to ensure that disturbances can not push the process towards the low value stable steady state, according to the domain of attraction. One choice is to select an operating point lower than the location of the second fold. A higher value can be acceptable but it depends on the disturbances (size and direction) .

The other approach is to control the process near the optimal operation point. The controller should be able to cope with disturbances which might push the process towards the wrong stable steady state in open loop. Conventional linear control will not be able to do the job, generally linear controllers will not be able to deal with this kind of nonlinear process behaviour, in satisfactory way.

Gain scheduled controller is considered to be implemented, where the gain should be proportional to the slope ($\frac{\partial Biomass}{\partial D}$) of the high producing steady state curve in the bifurcation diagram, meaning that the controller will have infinite gain at the bifurcation point, such that the bifurcation parameter, D will be reduced significantly. This approach has one obvious limitation, how to get the slope of the steady state curve.

Instead the gain scheduling controller is modified as follows:
One can line up the desired facilities of a controller:

- Safe operation at the optimal point.
- Rejection of disturbances which in open loop causes an undesired operation
- Elimination of the unstable steady state, if possible

To satisfy these requirements the controller should have sufficiently high gain around the optimal point to keep it there and away form the fold point. Therefore the gain should be increased around the optimal point and elsewhere it could be reduced to obtain better servo properties and to avoid the undesired properties of high gain proportional control, which would be the alternative. High gain proportional controllers tends to be oscillatory and actuator limitations is not easily accounted for in the design. It is also desirable that the control law is simple, such that implementation will not be a problem. Thus the idea is to construct a control law with low gain away from the set-point (the

optimal point) and a high gain around the set-point. The expression

$$\frac{x_{set}^p}{x_{set}^p + x^p} \tag{6.16}$$

will for low values of x be close to 1 and for large values it will approach 0, in both cases with a slope close to zero. At the set-point x_s the expression 6.16 is equal to $\frac{1}{2}$ with a gain of p. The proposed controller is :

$$u(t) = u_{set} + K\left(2\frac{x_{set}^p}{x_{set}^p + x(t)^p} - 1\right) \tag{6.17}$$

where $u(t)$ is the controller output and $x(t)$ is the state to be controlled. The parameters in 6.17 can be explained relatively easy, x_{set} is the desired set-point for the state x, u_{set} is the desired set-point for process input, K is a gain . The term $\left(2\frac{x_{set}^p}{x_{set}^p + x(t)^p} - 1\right)$ is a curve from 1 to -1 crossing zero at x_{set}, where the exponent p determines the slope around x_{set}. The parameters influence on the u-x relationship is illustrated in figure 6.2. How to select or determine the

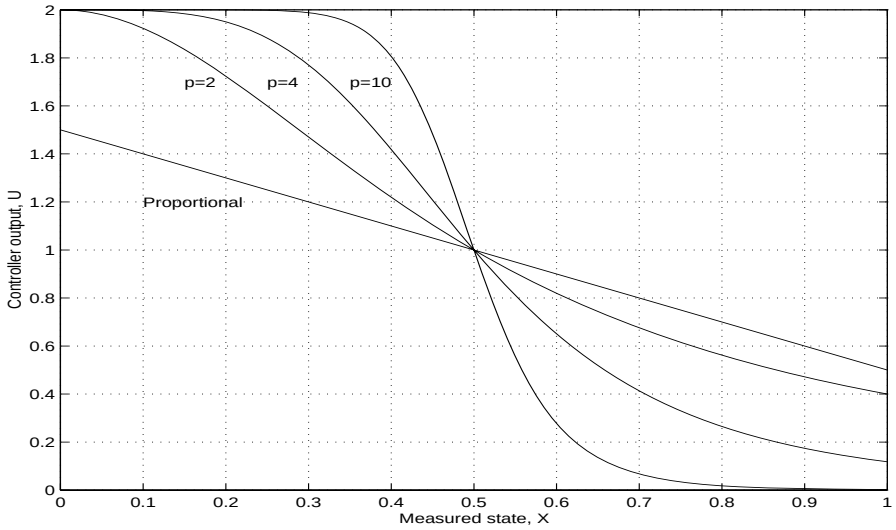

Figure 6.2: Comparison of controllers with different design parameters, $K = 1$, $x_{set} = 0.5$ and $u_{set} = 1$

parameters in a controller with as many parameters as this one is an obvious question, but the answer may partially be obtained from the bifurcation analysis and the optimization of process operation, as an initial estimate.

- x_{set} is taken as the optimal stable steady state,
- u_{set} is taken as the corresponding process input value,
- K should be large enough to pull the process back across the lower dilution rate with bifurcation, to return at the high producing steady state

close to the optimal operation point. As initial estimate one can select K as the distance from the optimal point to the bifurcation point located at the low production steady state, see figure 5.4.

- p will determine the slope in the u-x diagram at (x_{set}, u_{set}). The value of p is highly dependent on the process and the operation points distance from the bifurcation point. The closer the two points are located the higher a value of p should be selected to get a good performance.

6.6 Control of fold bifurcations

In this section a example will illustrate how the proposed controller work on a chemical reaction problem. The controller is able to globally stabilize the process by elimination of the fold bifurcation.

6.6.1 Auto catalytic CSTR

This example is one of the simplest chemical engineering models which exhibits fold bifurcations. The model Scott and Gray (1990) is given by a single differential equation.

$$\frac{dx}{dt} = \frac{1-x}{\tau} - x \cdot (1+\beta-x)^2 \tag{6.18}$$

The model is scaled with the feed concentration, x is the reactant, τ is the residence time in the reactor and β is the kinetic constant for the catalyst, in example this the value of β is set to $1/16$. The right hand side is divided into

$$L = \frac{1-x}{\tau} \tag{6.19}$$

and

$$R = x \cdot (1+\beta-x)^2 \tag{6.20}$$

R and L are plotted as a function of the conversion of reactant,$(1-x)$, in figure A.4. The bifurcation diagram can for this simple example be calculated either by hand or by numerical routines as CONT by Kubicek and Schreiber (1997). Figure 6.4 show the result of such a calculation. The bifurcation diagram is used for design of the controller.

Optimal production is calculated as,

$$\max_{\frac{1}{\tau}} \left(J = (1-x) \cdot \frac{1}{\tau} \right) \tag{6.21}$$

By isolation of $\frac{1}{\tau}$ from 6.18 at steady state one gets

$$\frac{1}{\tau} = \frac{x \cdot (1+\beta-x)^2}{1-x} \tag{6.22}$$

An alternative way to solve the optimal problem is so solve

$$\frac{dJ}{d\frac{1}{\tau}} = \frac{dJ}{dx} = \frac{d\left(x \cdot (1+\beta-x)^2\right)}{dx} = 0 \tag{6.23}$$

which yields the solution $x = 0.3541$ corresponding to a value of τ on 3.634. The productivity is plotted in figure 6.3. It is clearly seen that the optimum is located in the region where there exists multiple steady states with significantly different productivity.

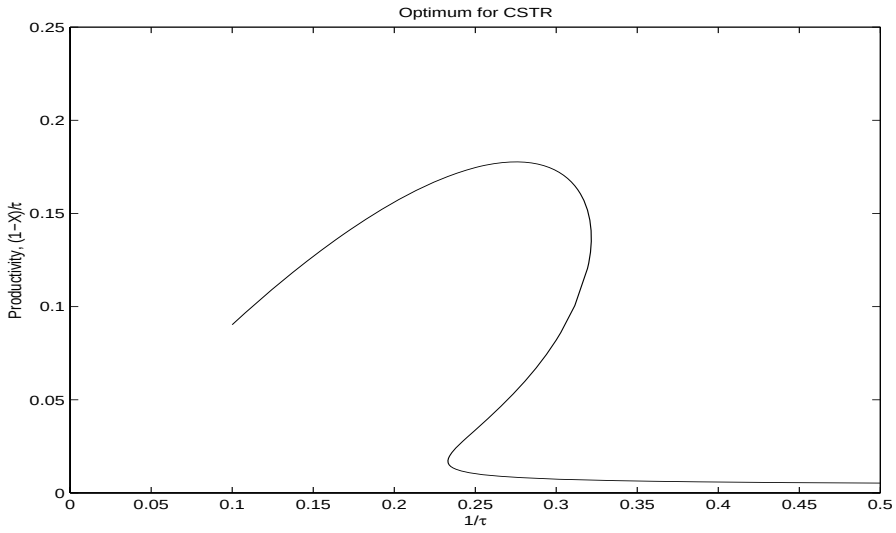

Figure 6.3: Productivity of auto catalytic reactor

Similarly to the yeast example shown in chapter 5 the middle solution is unstable resulting in a division of the phase plane into two areas divided by the unstable branch is shown in figure 6.4. Each area (known as domains of attraction) is characterized by the steady state solution which is attainable from an initial point located in the area. One part will be attracted towards the ignited steady state whereas the other part will be attracted towards the extinct steady state.

One way of illustrating this phenomenon is the effect of different initial conditions (initial conditions can be see as one type of disturbance). Open loop simulations show that the conversion is very sensitive to initial conditions. There is a clear correlation between the selection of τ and the initial condition where the separation occurs. Depending on which side of the unstable branch of the steady state curve in figure 6.4 the initial condition determines which stable steady state is reached finally.

If τ is set to the optimal productivity value one can see as in figure 6.5 how the state will diverge from the initial state and end up at either of the stable steady state.

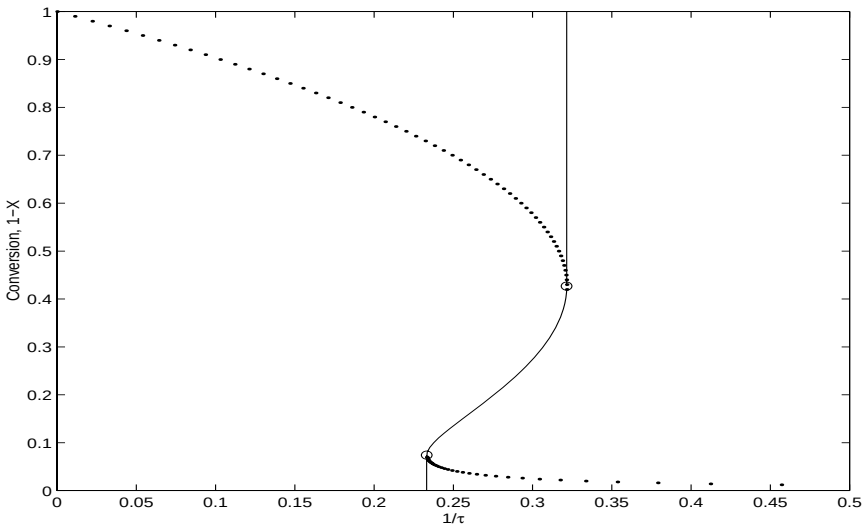

Figure 6.4: Domains of attractions for the auto catalytic reactor, the dottet lines indicate stable steady state, the full line indicates the separation of the domains of attraction. The circles are the fold bifurcation points. The full line between the bifurcation points is the unstable steady state.

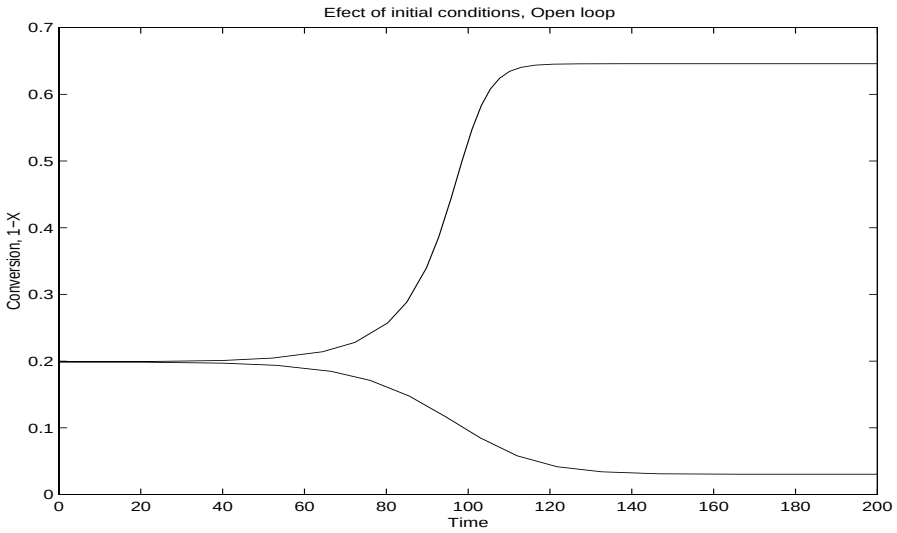

Figure 6.5: Simulation with different initial conditions at $\tau = 3.634$

6.6.2 Controller tuning

The control design problem is to assure operation at the optimal operation point and if possible eliminate the unstable steady state and also eliminating the effects of the initial conditions.

The distance between the optimal point and the bifurcation point is approximate $\delta\frac{1}{\tau} = 0.275 - 0.232 = 0.043$, this value is used as an initial guess for the controller gain, K. The power term p is set to 3. The result is shown in figure 6.6. As can be seen in the figure there are some problems at low conversions, so

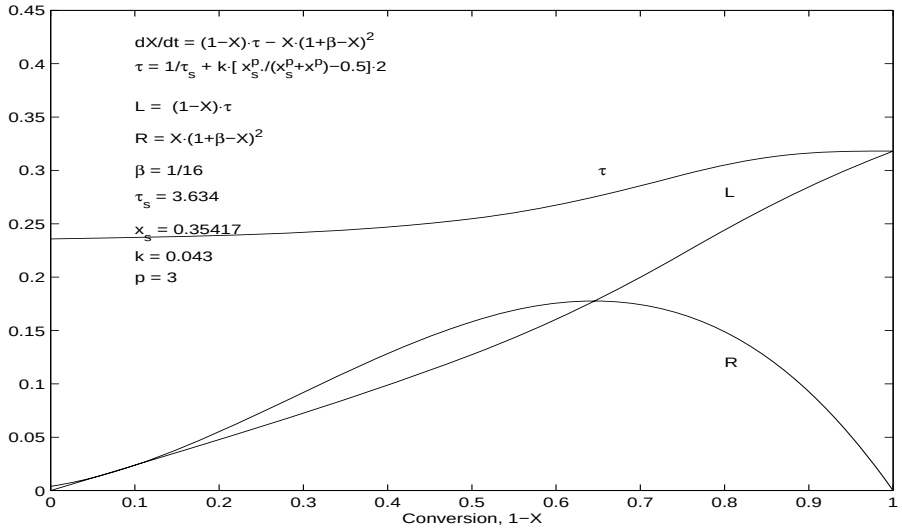

Figure 6.6: Closed loop coupling of R and L, there is an intersection of the curves at low conversion

there will still be a fold bifurcation. In order to overcome this problem one can increase the gain or increase the power term. In figure 6.7 the gain is increased and it is seen that there only exists one steady state for this set of controller parameters.

The simulation for closed loop operation with the selected parameters is performed and shown in figure 6.8. The closed loop simulations result in very good performance as was expected from figure 6.7 and 6.8, where there is only one steady state. This means that the sensitivity of the initial conditions is eliminated. The difference in the response times for the two closed loop situations can be explained from figure 6.7. The state derivative $\frac{dx}{dt}$ which determined the rate of change is determined as the vertical distance between the R and the L curve, as can be seen the distance of is large at high conversions whereas the distance is smaller at low conversions. Actually it is also possible to get slower response at high conversions if a control law is constructed such that the curve L is decreasing at high conversions thereby decreasing the rate of conversion.

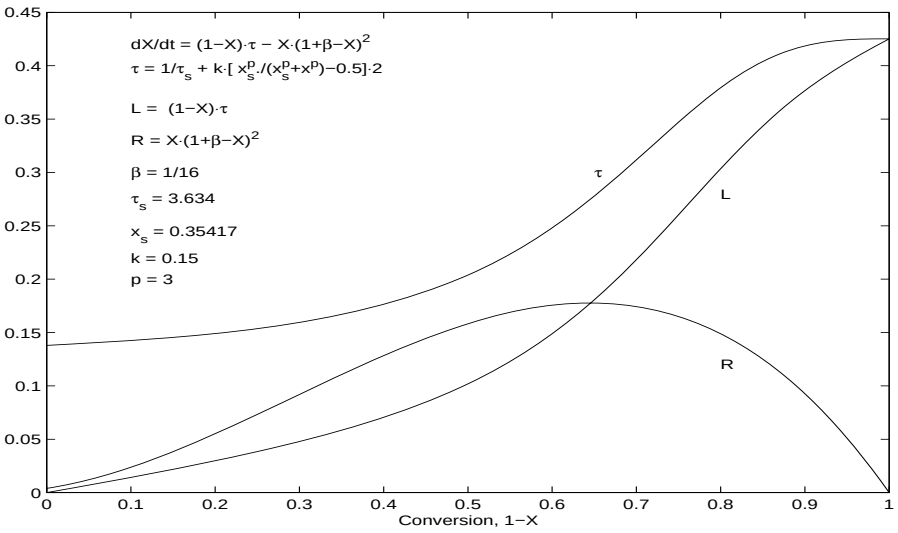

Figure 6.7: Closed loop coupling of R and L, controller is tuned with a higher gain, compared to the coupling in figure 6.6

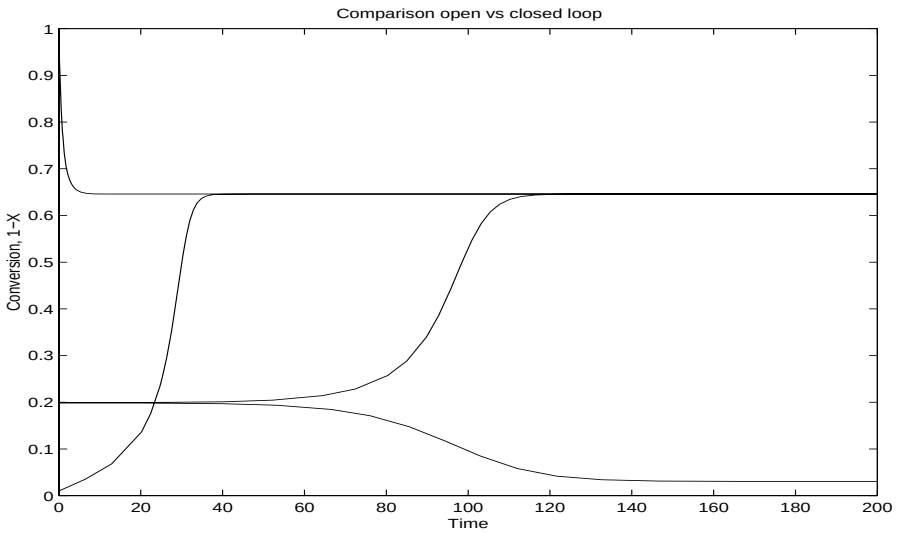

Figure 6.8: Simulation of Open and Closed Loop performance. The two slow responses correspond to the open loop simulations in figure 6.5 with slightly different initial conditions, the two fast responses with initial conditions $X_0 \approx [01]$ is operated in closed loop.

6.6.3 Conclusion on example

In this section the proposed controller is evaluated on a simple system. The simplicity of the system is used to analyze how the bifurcation arises in the process. By applying feed back it is possible to eliminate the fold bifurcation and thereby only have on stable operation point (the optimal one).

6.7 Control of Yeast cultivations

In this section the controller described in section 6.5 is applied to the continuous yeast cultivation with slight modifications.

First there is a discussion on which measurement are reasonable and which handles are available for control, also seen from an industrial point of view.

Measurements:

- **Biomass** is not appropriate as the controllable state since the level of biomass at optimal conditions is primarily determinated by the substrate feed concentration. The dynamics response of the biomass is relative slow.
- **Ethanol** is a suitable candidate but measurement might be troublesome at very low levels depending on the detection limit of the measurement device, typically measured by off-gas analysis. The model predicts a level around $0.01 mg/L$ at critical dilution rate.
- **DOT** (dissolved oxygen tension) is measured by electrode which will provide information on oxygen consumption (q_{O_2}).
- **Carbon dioxide** is measured by electrode, leads to CO_2 production q_{CO_2}.
- **Temperature** Measured by electrodes.
- **pH** Measured by electrode

Control arising:

- **Dilution rate** This is an obvious handle for control since the feed flow in and out of the reactor can easily be manipulated by control valves.
- **Feed Concentration** Feed concentration can only be diluted, can be achieved by ratio control.
- **Cooling jacket flow rate** Can be manipulated by control valves.
- **Base and acid addition** Flow addition by control valves.
- **Stirrer speed** By manipulating the power supply to the motor the rpm is adjusted.

- **Air flow rate** By manipulating the power supply to the pump volumetric flows is changed.

It is assumed that the reactor temperature is controlled by adjusting the flow rate through the cooling jacket. Similar is the pH controlled by the addition of acid and base, depending on the operation point according to some control law. The DOT is controlled by the stirre speed or/and the air flow rate. These loops is considered to be operating properly such that they will not influence the operation of the reactor.

Biomass concentration X is assumed to be measured either directly or estimated by an observer.

The measured value is selected as the CO_2 production, which is used for control at the high biomass producing steady state since the effect of a change of metabolism will immediately show in the CO_2 production, since much more CO_2 is produced by the overflow reactions.

Similar to the biomass level the CO_2 level is at the desired operation point determined by the available amount of substrate.

To take advantage of this the following process measurement is suggested

$$CER = \frac{q_{CO_2}}{X} \tag{6.24}$$

where X is the biomass concentration. The CER is closely related to the biomass formation rate and at steady state the value of is constant and independent of the dilution rate as long there is no ethanol production. Ethanol contributes to the CER as well and that is what we want to take advantage of. The advantage of using this combined measurement is that if the overflow metabolism is triggered the it will show immediately in the q_{CO_2} measurement and action can be taken. If the cause is an increased substrate supply then the biomass concentration will increase slowly.

As control handle is chosen the dilution rate since it is easily adjusted and the bifurcation analysis in chapter 5 showed that the substrate feed concentration did not affect the location of the bifurcation point in the (S_f, D) diagram, see figure 5.7.

The purpose of control is to keep the yeast cultivation at the optimal point and be able to reject the disturbances that might occur .

The evaluation of the proposed control scheme is done by simulation of the closed loop system illustrated in figure 6.9 which show that the controller is performing very well , even with large disturbance in the feed concentration (e.g from 30to50g/L).

6.7.1 Evaluation of closed loop continuous yeast cultivation

The general idea of the controller is described in the previous section is adopted. The result was simulated with a sample time of 15 min. and a time delay of the measurement of the biomass of 1 hour. The result of the disturbance simulation is given in figure 6.9

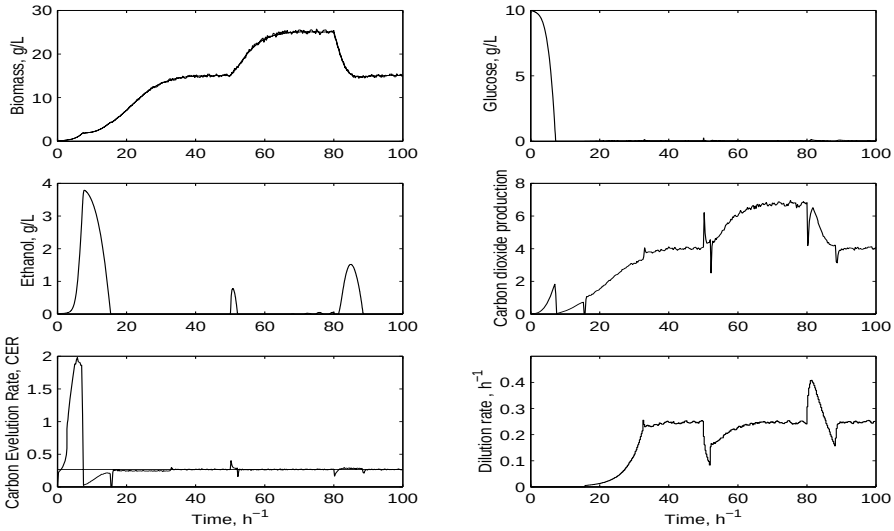

Figure 6.9: The result of cultivation startup and disturbance simulation with disturbances in the feed concentration. After steady state has been reached, $t = 50h$ the feed conc. is changed from $30g/L$ to $50g/L$, at $t = 80h$, feed conc. is reset to $30g/L$. The set point of the measurement is $CER = 0.26$, $D_{set} = 0.24946$, $K = 0.05$ and $p = 3$

The figure illustrate that the controller is able to recover from the disturbances and return to the desired steady state. Even the delay in the biomass measurement does not cause a problem, as can be seen in the CER measurement. At the step-up the ethanol production is captured relatively fast.

Even with these large disturbance and the biomass measurement delay and the slow sampling it is possible to maintain operation at the desired operation point, even when the disturbances initiate ethanol production the closed loop process is able to recover.

6.7.2 Control of fed-batch yeast cultivation

Again it is desired to control the reactor such that the glucose is directed into biomass and not into ethanol. In the literature it is widely accepted that in fed-batch the feed profile should be exponentially increasing substrate as long as no other constraint is active (limitations in oxygen transfer and heat transfer are typical constrains). To determine the exact feeding profile one has to apply dynamic optimization routines which is nontrivial to apply to such a stiff system as the yeast cultivation, with large variations in time constants. To get an idea of the complexity of this type of optimization consider the different time scales of the states glucose is directly influenced by the feed rate whereas the biomass is responding slowly to feed changes. To solve the dynamic optimization problem the collocation should be such that the fast dynamics is captured, this will result in a very high number of collocation points. Otherwise the fast states

should be considered to be in quasi steady state and the differential equation is transformed to an algebraic one. In order to avoid dynamical optimization we considered to apply the developed controller for the fed-batch case since it was able to maintain operation at the optimal productivity, which for the simple case corresponded to the maximal growth rate.

With a minor modification the control algorithm looks as follows:

- Measure qCO_2 and estimate the biomass
- Calculate the specific CER
- Calculate the controller action $F_{in} = F_{set} + K\left(2\frac{CER_{set}^p}{CER_{set}^p + CER(t)^p} - 1\right)$
- Set $F_{set} = F_{in}$
- Set $K = F_{set}/2$ if $F_{set} \geq 2$

The last two statements are included to increase the flexibility and having the possibility to generate the exponential feeding which is desired for the fed-batch case.

With noise on the measurement of CER and biomass estimation 2 and 5 % respectively, simulation of the fed-batch with the above controller is performed. The result is shown in figure 6.10 and is very promising. The exponential feeding profile is generated automatically without previous analysis. The noisy signal for glucose is due to the constant feeding which is changed in steps. The

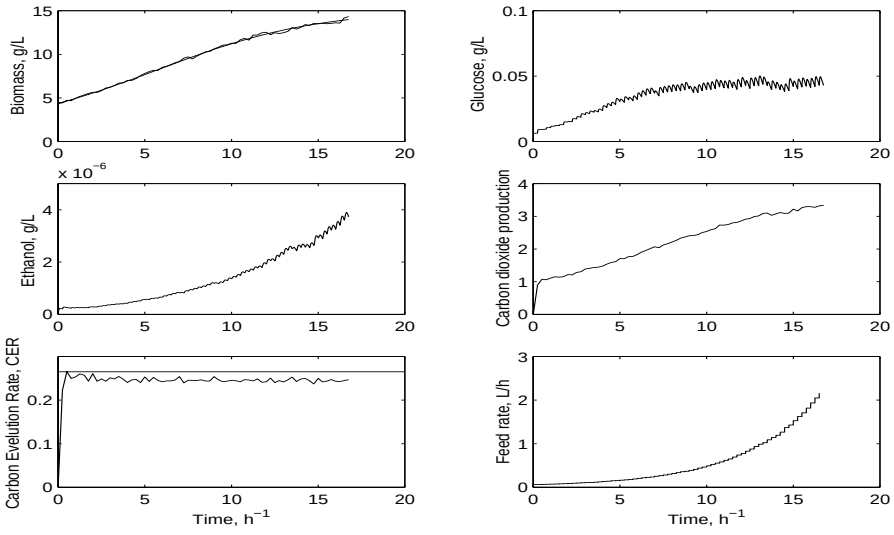

Figure 6.10: Simulation of Closed loop fed-batch Yeast cultivation. Simulation is terminated when the volume is above 10 L.

simulation show that the ethanol level is increasing but at a very low level, e.i $5\mu g$, which is desired. The glucose level increases to around $50mg/L$. The reason for the noisy glucose and ethanol is that the feed rate is increased step

wise at each sample time. The level of q_{CO_2} is increasing due to the increasing amount of biomass. The limit for the simulation is set to the maximum volume of the reactor $V_{max} = 10L$.

6.8 Conclusion

This work has shown how a new type of controller can be developed for fold bifurcation control and as seen in the first example the controller is able to eliminate the occurrence of the fold and ensure operation at the optimal point.

The tuning is straight forward from the bifurcation analysis and the parameters are related to the physical variables. In the second example the controller proved to be powerfull for disturbance rejection. The step up simulation of continuous cultivation showed that the controller can cope relatively large disturbances in the feed concentration and eliminate the production of ethanol.

The modified control algorithm has shown to be able to generate exponential feed profiles, without prior specifications.

The controller have been applied in experimental setting Lei (2000). There it is compared to a incremental PI and a LQ controller, with a slight modification of the controller it is recommend over the PI and LQ controller.

7

Optimization of Penicillin Cultivation.

Dynamic optimization of fed-batch process is a nontrivial problem, especially if there are large differences in time constants. Typically substrate have relative fast dynamics since it is directly influenced by the feed (in the range of seconds), whereas the biotic components have slow dynamics (in the range of minutes). The result of the optimization reveals an unrealistic result. The reason for the strange result is that by applying optimization the model is being pushed beyond its limit of applicability.

7.1 Introduction

Fed-batch operation is widely used in the chemical and biochemical industry. The industrial desire for optimal operation of fed-batch operation poses the problem of optimization of a time varying (dynamic) process, which is interesting from an academic point of view.

The method of Biegler (1984) has been applied to a morphologically structured model for a filamentous fungus describing the growth of biomass and Penicillin production, developed by Zangirolami et al. (1997). The resulting optimal feeding strategy is at first surprising but we believe that it can be explained with help of the morphological model.

First the principles of the collocation are demonstrated and the optimization problem is formulated. The process model of the penicillin producing organism are presented and the objective criterion and the constrains are given for the optimization. Finally the result of the optimization is discussed and conclusion is drawn on the basis of the results.

7.2 Dynamic optimization of fed-batch fermentation

In this is presented a method for solving dynamic optimization problems and show how it can be applied to the optimization of a fed-batch cultivation of a filamentous fungi producing Penicillin. The basic problem in dynamical optimization problems is that the differential equations, which describe the process

behaviour, has to be solved in some way at each iteration step of the optimization. Direct simulation will be rather time consuming compared to the optimization step. Another problem is how to determine the gradients, needed for the optimization routine, from the simulation. The approach in this work is to apply a time discretization of the differential equations and thereby transforming the dynamic problem into a formulation where the differential equations are replaced by a set of algebraic equality constraints. The advantage of this approach is that gradients are easily obtainable from the discreticed model and that the solution of the process equations and optimization are performed simultaneously. The method used for the discretization of the differential equation is orthogonal collocation, which is a polynomial approximation of the states.

For further details on the method of collocation see section 6.3.2, where the discretization and collocation method is described.

7.3 Formulation of the Optimization problem.

In dynamical optimization the problem can be formulated as; find the conditions such that a given objective is maximized (minimized) according to the differential equations and other constraints. In mathematical terms this can be expressed as

$$\min_{u,x_0,T_f} \Phi(x,u,t) \tag{7.1}$$

such that

$$\frac{dx}{dt} = F(x,u,t)$$

$$x_l \leq x \leq x_u$$

$$u_l \leq u \leq u_u$$

In addition to these constraints there can be both soft and/or hard constraint equations, typically given by

$$g[x,u,t] \geq 0$$

$$h[x,u,t] = 0$$

In the literature this type of problem has been addressed by a number of different approaches

- Analytic solution, e.g variational calculus Pontryagin et al. (1964).
- Iterated Dynamical programming, e.g Luus (1993)
- Simultaneous approach, e.g Biegler (1984).

These approaches have been applied to a large number of different dynamical optimization problems. The different approaches have different strong and

weak points. A major drawback of the analytical approach is that it is only feasible for small size problems. The idea of applying orthogonal collocation was introduced by Biegler (1984). Applying the method of orthogonal collocation described previously, the differential equations will be reformulated into a set of equality constraints (the residual). Now the problem is in a form which is solvable by standard Non Linear Programming algorithms such as Successive Quadratic Programming. Variables for the optimization are all the parameters x_i and the input profile. In this work the input profile will be given as a polynomial with fixed order. In an implementation situation the polynomial should be replaced by a simpler approximation of the input profile. The major advantage of applying the polynomial approximation is that the extraction of gradients becomes very easy and straightforward, and the use of Lagrange polynomials ensures that the polynomial coefficients have a physical meaning, i.e. the state value at the collocation point.

7.4 Description of the Penicillin model.

The model described in this section is similar to the on presented by Zangirolami et al. (1997) The only change is that the microorganism is allowed to grow on Corn Step Liquor (CSL) until that has consumed. In the following both growth and production are based on glucose as the limiting substrate. The model is morphologically structured, which means that the model include the dynamics for the morphology of the biomass. The morphological states are the apical (the tip ends of the fungi), the subapical (the region just behind the apical compartment) and the hyphal compartment (the main structure of the branched fungus) as illustrated in figure 7.1. The model describes the consumption of substrate (glucose) and production of penicillin. The reason for developing a morphologically structured model is that it has been shown that the production of penicillin only takes place in a part of the biomass, in the subapical and a part of the hyphal compartment but there is no production in the apical compartment. The model is relatively simple with only seven states and it has only 13 parameters.

7.4.1 Model Equation and Parameters.

The cultivation reactor is operated in fed-batch mode, but modeled with a continuous sample withdrawal rate of $0.029 l/h$.

The penicillin model is described by the following set of equations:

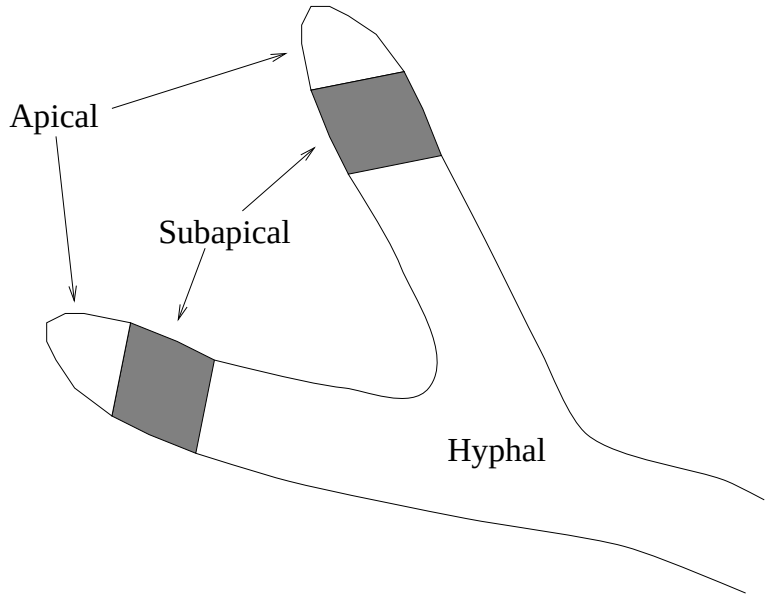

Figure 7.1: Morphological structure of the Penicillin fungus

Reaction rates

$$\begin{aligned} r_1 &= \tfrac{k\cdot x_5}{kcs+x_5} \\ r_2 &= r_1\cdot(x_2+x_3+fh\cdot x_4) \\ r_3 &= ku1\cdot x_3 \\ r_4 &= ku2\cdot x_2 \\ r_5 &= \tfrac{ku3\cdot x_3}{(x_5\cdot Kuc3+1)} \\ r_6 &= \tfrac{k2\cdot x_5}{(x_5+kc2+x_5^2/Ki)}\cdot(x_3+fh\cdot x_4) \\ D &= u/x_7 \end{aligned} \tag{7.2}$$

Dynamical equations

$$\begin{aligned} \tfrac{dx_1}{dt} &= (r_2-D+0.0229/x_7)\cdot x_1 \\ \tfrac{dx_2}{dt} &= r_3-r_4+x_2\cdot(r_1-r_2) \\ \tfrac{dx_3}{dt} &= r_4-r_3-r_5+x_3\cdot(r_1-r_2) \\ \tfrac{dx_4}{dt} &= r_5+x_4\cdot(fh\cdot r_1-r_2) \\ \tfrac{dx_5}{dt} &= D\cdot(s_f-x_5)-(\alpha\cdot r_2+0.6878\cdot r_6+ms)\cdot x_1 \\ \tfrac{dx_6}{dt} &= x_1\cdot r_6-D\cdot x_6 \\ \tfrac{dx_7}{dt} &= u-0.0229 \end{aligned} \tag{7.3}$$

The states are as follows:
x_1 : Biomass [g/l]
x_2 : Apical compartment
x_3 : Subapical compartment
x_4 : Hyphal compartment
x_5 : Glucose [g/l]

x_6 : Penicillin [g/l]

x_7 : Volume [l]

The parameters are given as:

$s_f = 450$, $k = 0.14$, $kcs = 0.0015$, $k2 = 1.3541$, $kc2 = 0.0132$, $Ki = 0.0101$,
$ku1 = 2.3$, $ku2 = 0.7$, $ku3 = 0.19$, $Kuc3 = 20$, $fh = 0.13$, $\alpha = 2.2$, $ms = 0.0281$,

7.4.2 Optimization Criterion and Constraints.

A number of constraints has to be imposed on the system, due to the physical equipment . e.g. the upper limit on the **volume**, the **compartments** is defined as fractions of total biomass, which leads to upper and lower bounds, **glucose** is bound by the feed concentration. Due to modelling assumptions, e.g an upper limit of the **biomass concentration**, in order to avoid oxygen limitations. If this limit is exceeded the fungus tends to floculate and form pellets, which will lead to a lowered production of penicillin due to the introduced diffusion problems and the model will no longer be valid.

Another constraint imposed due to modelling limitations is an upper limit on the cultivation time. If the cultivation runs for a longer period of time mutations might dominate and the behaviour of the microorganism can not be predicted by the model. For further reading about the model consult the original paper.

The above mentioned constraints has to be included in the optimization as hard constraints, which are not allowed to be violated. The constraints are

$0 \leq$ Biomass $\leq 50 g/L$

$0 \leq$ Apical compartment ≤ 1

$0 \leq$ Subapical compartment ≤ 1

$0 \leq$ Hyphal compartment ≤ 1

$0 \leq$ Glucose $\leq S_f$

$0 \leq$ Penicillin

$0 \leq$ Volume $\leq 35L$

Final time $= 200h$

The state constraints above can be implemented in most optimization algorithms a they define the space (in mathematical sense) within which the state variables can vary according to the differential equations in section 7.4.1.

7.4.3 Objective function for Penicillin production.

An important factor in the optimization of any process is the objective function $\Phi(x,u,t)$. The selection of the type of objective naturally depends of the process. In the case of penicillin production the objective could be formulated as: Produce as much penicillin in the shortest time, in as high concentration as possible with use of a minimum of glucose. One way to turn this into mathe-

matical terms could be to set $\Phi(x,u,t)$ to a combination of the following terms

$$\Phi_1 = \frac{P(T_f)\cdot V(T_f)}{T_f} \tag{7.4}$$

$$\Phi_2 = P(T_f) \tag{7.5}$$

$$\Phi_3 = \frac{P(T_f)\cdot V(T_f)}{\int_0^{T_f} S_f \cdot u(t)dt} \tag{7.6}$$

The first term expresses the production per time (important for the turnover of the process), the second term the end concentration (important for downstream processing) and the third term expresses the utilization of the glucose (important cost factor for the substrate). A linear combination of the three terms is

$$\Phi_{Total} = \sum_{i=1}^{3} C_i \Phi_i \tag{7.7}$$

The problem with this objective is how to determine the constants, C_1, C_2, C_3, which is important due to the different scales of the three individual objectives. Cost estimates would be desirable, but in general these cost estimates are not constants due to changing market situations.

The initial investigations showed however that the resulting profile was not changing significantly as the objective function weighting was changed. This phenomenon can be explained by the fact that the three individual objectives basically are non conflicting, due to the bounds on the final volume and the final time. Therefore the simple objective function used in the following is

$$\Phi(x,u,t) = -P(T_f)\cdot V(T_f) \tag{7.8}$$

The objective is the same as Φ_1 but the direct influence of the final time is not omitted since the final time is fixed. Simulation studies show that the glucose concentration responds very dramatically to changes in the feed rate. Compared to the other states it reacts much faster. This point has to be accounted for in the study of the relationship with fast and slow dynamics.

The collocation method will require a high order polynomial to describe the fast dynamics, whereas the slow dynamics only require a lower number of collocation points. This problem is solved by assuming that the fast dynamics is at quasi steady state. This corresponds to setting the differential part of that residual equal to zero. The result of this change is that the number of collocation points can be reduced significantly, this speeds up the optimization, without decreasing the accuracy of the solution.

7.5 Optimization Results.

The original model by Zangirolami et al. (1997) contained a term describing the growth on complex medium (corn step liquor, csl). The operation was

initiated with a batch like phase where the growth of biomass was on basis of the csl, but due to the maintenance term in the model a small addition of glucose was needed in the batch phase. After the batch phase the growth and production (fed-batch operation) was on the basis of glucose and no csl was present in the cultivation broth. To improve the optimization the problem is split up into a batch (growth) phase and a fed-batch (production) phase.

7.5.1 Growth phase optimization

The growth phase is given a set of fixed initial conditions and the objective is to perform the batch in the shortest time possible. This means that the batch time is mainly determined by the initial concentration of the csl. The problem is to determine the addition of glucose needed for maintenance.

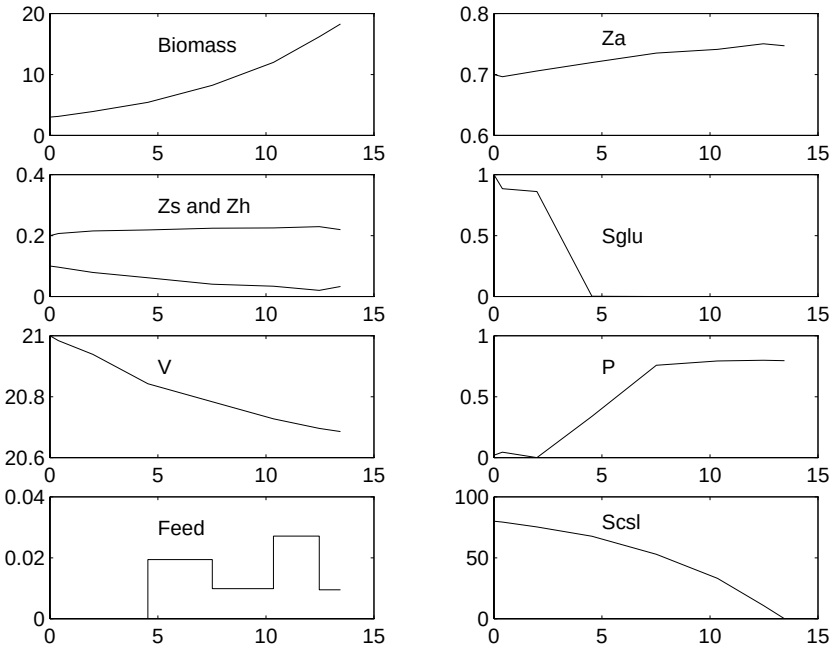

Figure 7.2: Result of the Optimization of the Penicillin model in the batch phase

Figure 7.2 shows that the feed of glucose is zero until the glucose in the reactor is depleted and the feed is turned on to balance the need of the microorganism. The feed is piecewise constant within the collocation interval, for several reasons: It is the simplest formulation (only one parameter) which reduces the dimensionality of the optimization problem. Usage of a higher order polynomial would only have a small impact on the result and it is implementable in a real cultivation system. The result is that biomass is growing exponential on the csl and there is only small changes in the structural compartments. Note that the small decrease of the reactor volume is not an error but is due to the continuous sample withdrawal from the broth included in the model.

7.5.2 Production phase optimization

Since the csl has been depleted when the production phase begins this state is removed from the model to simplify the next optimization problem. The initial state vector is otherwise given by the result of the batch phase optimization. Again the simple zero order hold type of feeding is applied. The result of the production phase optimization is shown in figure 7.3. The feed rate is kept low

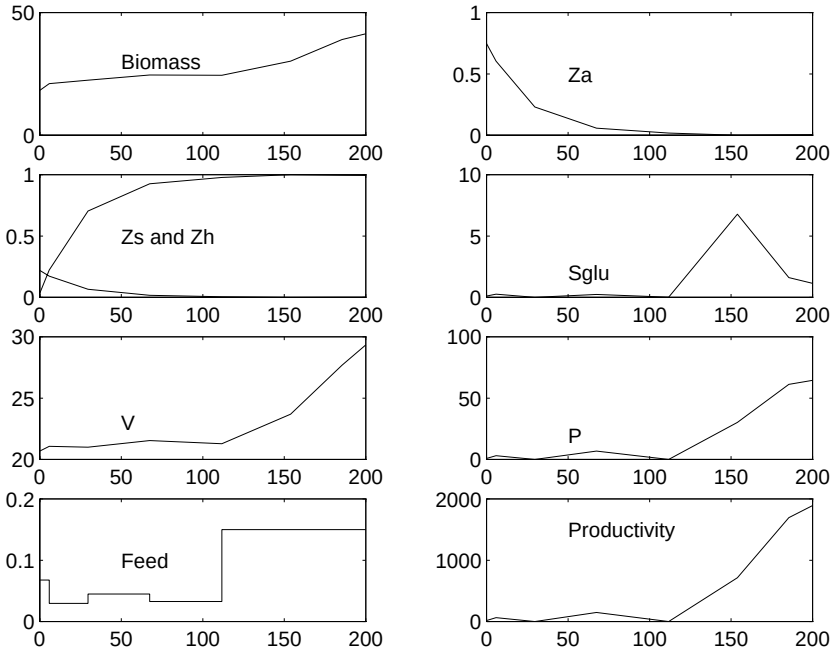

Figure 7.3: Result of the Optimization of the Penicillin model in the production phase

for the first 110 hours and then the feed rate is increased and kept constant for the rest of the cultivation period. The result of this rather surprising feeding strategy is that the biomass concentration is relatively constant during the first period. What happens during this long phase is a transition in the morphological structure of the biomass, such that the hyphal compartment becomes dominant. When the feed is turned up the production of penicillin is initiated together with growth of the biomass. The result of the feeding profile is that the biomass concentration ends up at its upper limit $X_{biomass}(t_f) = 50$ g/L. In model term this can be explained by the fact that the biomass growth gets contributions from the apical, subapical and a part of the hyphal compartment whereas the penicillin production takes place in the subapical and a part of the hyphal compartment. When the biomass composition is mainly hyphal compartment, growth and production both benefit. This means that the first part of the production phase is used for minimization of the apical compartment such that growth and production is balanced. Again as for the batch phase a slight improvement of the productivity could be obtained if a higher order polynomial is applied as a feed rate profile, but the complexity of the optimization is

increased and the result is not significantly different from that shown in figure 7.3

7.6 Conclusion.

This study of the optimization of the morphologically structured model for penicillin production have shown that the coupling between growth and production is very essential, in this case the coupling lies in the different biomass compartments. The problem with the results can be that the model is not valid for a feeding strategy of this type. The optimized result is quite different from the conditions where the model was developed and validated. This case show a possible use of the dynamical optimization in the field of model validation, if the optimization problem is realistic then the optimization will drive the model out into some extreme area of the optimization space.

The optimization result can not be expected to be directly translated into an experiment with a similar result, but can be used as guideline for planning of experimental work, such that the model can be improved, to become valid over a wider range of the state space.

The focus should be on including regulatory mechanism which describes how the genetic aspects is expressed in process behaviour. This will improve the prediction power of the model and therefore a more reliable result will be the outcome of such an exercise, this is clearly an iterative process.

8

Conclusion

Here is the main conclusion of the thesis summarised. The scope of the project was to investigate the influence of nonlinear phenomenon on the operation of fermentation processes. Special attention on the phenomena around the critical dilution rate is at focus.

The development of models describing biochemical phenomena is an interesting field of research. There are many interesting details which can cause a dramatic effect on the state of the microorganism. The understanding of these details are valuable in the understanding of biological phenomenon such as the problems with simultaneous uptake of substrates e.g. chapter 2, the switch in metabolism treated in chapter 3. The strength of mathematical models is the amount of information that they can contain and the flexibility they poses. Depending on the area of application different levels of detail is necessary (and sufficient), these points are discussed in chapter 2.

The finer details of switches and triggers is on an operational basis less relevant since the control variables are not related directly to these trigger mechanisms. The understanding of cell metabolism is central in the selection of the components to model. The model described in chapter 3 is intended to be used for detailed analysis of yeast fermentation with special attention on the shift in metabolism around critical dilution rate. The model has many of the desired properties, physical transparency, flexible to operation conditions. The relative simple model is able to describe with sufficient accuracy batch, fed-batch and continuous fermentation. The components are selected to reflect the physical behaviour to give meaning full interpretation of results, even though there are many simplifications in the model development.

The result of the 'real life' data analysis was presented in chapter 4 where a set of experiments is analysed to investigate the how closed loop data can be used in model identification. The technique is a bit more complex since the identification is divided into a two step identification problems. The result of the two step identification procedure is a better estimate of the process model at the operation point. The procedure have showed to improve the result when compared with a straight one shot estimation based on process input/output data. These results contribute to the recommendation of the two step identification procedure. It should be kept in mind that some of the operation points are unstable and not attainable in an open loop experiment.

The estimation results obtained in the region around the critical dilution rate,

where there should be multiple steady state, indicate that the planning of the experiments should be paid more attention than is the case with the available set of data. The result obtained by applying direct identification on input/output data indicate that there are possible instability in the process, but the results are not rather uncertain. It is evident that the process undergo a dynamic change. The dominating time constant is changing from operation point to operation point. This is a clear evident for nonlinearities.

The main contribution in this work is the analysis of the yeast fermentation presented in appendix A and chapter 5. The analysis showed that the model developed in chapter 3 based on physiological insight as presented in chapter 2 contains multiple steady states in a region around the critical dilution rate. The meaning of the multiple steady states is the operational difficulties caused by the relative larger domain of attraction of the low biomass yielding stable steady state. If the conditions at the operation point is pushed slightly to trigger the formation of acetaldehyde and ethanol, the only way to retrieve the high biomass yielding stable steady state is to reduce the substrate feeding rate significantly to avoid additional ethanol production. The confusion of the microorganism is obvious, since the production of ethanol can be viewed as a feast mechanism. The production of ethanol is a mechanism to cope with the surplus of energy which is taken up from media. The surplus of substrate trigger the increased formation of proteins for this production. If then the energy source is depleted by a decrease of the feed rate, the conditions seen by the microorganisms is changed to more famine conditions where ethanol is the major energy source. In the extreme case the feed is turned off. Then the situation is as for the batch case and lag phase would bee seen. The feed of substrate should be set back to the original settings when ethanol i depleted.

In chapter 6 the intention is develop a control strategy to handle the fold bifurcation described in the previous chapters taking the above considerations into account. The result is a controller which acts very strongly to hold the process at the desired operation point. This controller is tested in simulation with reasonable disturbances for revealing the effectiveness of the scheme. The result of these simulations show the controller is very effective in controlling the continuous fermentation process. The control scheme is modified slightly and the controller can be applied for solving the fed-batch fermentation problem. Compared to the expected result of a optimization of the feeding profile (exponential growth $\Rightarrow$ exponential feeding) the control performs very well. This type of analysis of biological models of this complexity is new and the results is in good agreement with what is obtained for simpler models, this could indicate that the predictive strength of the model is rather good.

The operation the fermentation process in fed-batch is treated in chapter 7. The model used as a case is quite different form the yeast model. The technique which are applied to the penicillin fermentation is dynamic optimization. This method is relative well tested for simple biological models and classical chemical process. The result of the optimization is revealing the danger of applying models to conditions where the prediction is relatively poor. The real test of the

optimization is to apply the generated profile to a real fermentation, but here the result showed a weak spot of the model.

The major result of the process model analysis is a strong evidence of the multiple steady states and straight forward ways to overcome the difficulties that arise from the fold bifurcation and the unstable steady state.

The unstable steady state is a strange invisible academic state.

- It does not exist in operation (by nature unstable)
- For systems larger than one or two states the interpretation of a border line between the stable states disappears since the states might spin around in the space spun by the full size operation space, determined by the domain of attraction for the individual stable steady states.

The unstable steady state can be stabilised by control an tracked through domain as in Møller (1993). Another example is the one dimensional CSTR case with control of conversion where it is possible to track the reaction rate curve.

The outset of the work was to investigate how information on the cells genetics can be used in the operation of fermentation processes. The model described in chapter 3 the genetic information is translated into mathematics and is represented in the rate equations as a number of inhibition and activation terms.

The analysis of process models have shown that the existence of fold bifurcations and multiple steady states in fermentation process is existing. This is supported both by the model analysis and the data analysis. As mentioned experimental studies and theoretical work have also pointed this out. In this thesis the size of the process model is large compared to other studies.

9

Future Directions

The work presented in this thesis have given some tools and guidelines of applications of these tools. The integration of the tools in the design and operation in the biotechnological process industry is still waiting to be seen. A number of steps on the way are presented to provide directions which might lead to application of system engineering tools in industry.

9.1 Introduction

The result of the thesis is discussed in chapter 8, but here the concern is what is needed to carry the ideas in this work into the industry. This are based on this work and the provided insight; what is the perspectives?.

The immense amount of information on microorganisms such as *Saccharomyces cerevisiae*, has to be structured in some way to gain the full benefit from this knowledge. The way to structure this knowledge is by using mathematical models. The model presented in this thesis is to be viewed as a starting point for further modelling studies. This will be discussed in section 9.2. More levels of modelling are to be include genetic information into the modelling. There have been some initial attempts in the literature on how the structure of such models should be.

Another approach to structuring the large amount of information held by process data, is to apply the ideas in chapter 4. The can be applied to the process data and useful information can be compiled in this way for the production, since the experiments can be performed at normal/optimal operation conditions, with the appropriate input signals it should be possible to find suitable transfer function models. The information gained by this could be used for better control of the process at these operation conditions.

The analysis of the developed model showed that the characteristics of the simple model with growth inhibited kinetics is preserved. The analysis should be performed in order to reveal any undesired modes of operation. Therefore it is strongly recommended that the models which is used for designing a given process, this can be in the area of biotechnology or in any other field of engineering, is analysed in this way. The results obtained in this showed how the problem get worsened by the increase in substrate concentration. Another feature that could be investigated is if the assumption on ideal reactor conditions does not hold, how will this affect the analysis. There could be oscillations in

the system due to the gradients of the certain compounds.

9.2 Model developments

The model presented in chapter 3 will provide a sound basis for further model investigations. The model can be used to design detailed experiments to reveal the some of the underlying mechanisms which regulate the behaviour seen in yeast fermentation. The focus of the model development is as mentioned in the initial parts of this thesis is depending on the application of the model.

One interesting experiment could be to obtain measurement of the two active biomass components, this is related to development of sampling techniques and methods of analysis of these components.

The demands from industry to apply models in both process and control design, are easy to use and understandable models. The models should be flexible in the sense that if a new strain is developed/selected in lab then the effort to adapt the model of the microorganism should be reasonable.

The points mentioned above is in a way looking into the microorganism, to cover some of the assumptions mention in chapter 3. Another perspective is to view the surroundings of the microorganism and develop the model of the tank reactor to include the gradients which are a complicating issue in industrial scale due to the size of equipment. The assumption of ideal mixing is not valid in that size. There will be gradients of a number of key components and the mechanical design of the vessel and the impeller will be of importance. Mixing properties of oxygen and substrate is depending on the viscosity of the broth which again can be affected by the density of the biomass. One approach is to divide the vessel into different zones, e.g. close to feeding points, around impeller, between impellers, around impeller axle, close to wall, top and bottom.

9.3 Development of tools

The tools applied in this thesis are at present not well suited for implementation in the process industry. The nonlinear model predictive optimization are at present to slow to deal with any problems of a realistic size, this will be overcome by an increase of computational power and development of faster routines. The identification methods presented here is used only in off-line analysis but the ideas should be reformulated into an adaptive scheme to improve the model identification. The result should be improved regulatory properties. The method of obtaining process data for model analysis in closed loop will help to provide new insight to the occurrence of unstable processes.

A

Bifurcation theory

The science of bifurcation analysis is a well established area in the mathematical community and results are beginning to emerge from other areas such as chemical reactions systems and chemical processes. The intention with this appendix is to give an introduction to the field in the light of the application presented in chapter 5. What is bifurcation analysis ? In general terms bifurcation analysis is the an investigation of the stability properties of a system.

Science is the knowledge of consequences, and dependence of one fact upon another. Thomas Hobbes (1588-1679)

A.1 Introduction

In this appendix the concepts in the area of bifurcation theory are introduced and discussed in relation to the application in the later chapters in this thesis. The area of bifurcation theory is a broad field of mathematical analysis of models/systems with respect to dynamical behaviour and stability of solutions. The field is treated in standard text books such as Thompson and Stewart (1986). For more mathematical description of the phenomenon a reference is Guckenheim and Holmes (1983). The phenomena which are dealt with by this analysis are all nonlinear phenomenons which is not seen in linear models where the principle of superposition holds. The structure of the chapter is as follows, first the basics and terminology are presented, then two simple examples are introduced to illustrate the concepts and finally a short presentation is given of the numerical algorithms.

A.2 Mathematical setting

The systems of interest in this section are nonlinear processes which may have some interesting properties from a mathematical point of view. These properties can cause operational problems of the process. The mathematical properties are only existing for the nonlinear description of the system and not for a linear process description, which is typically applied for process control and optimization. Whereas the analysis is carried out using linear theory. The analysis is an iterative procedure where the set of solutions in the nonlinear world is

analysed locally, with linear methods, and these local results are then patched together to obtain the global result.

The mathematical basis of the analysis is the model description in form of a set of ordinary differential equations

$$\frac{dx}{dt} = f(x,t,\phi) \tag{A.1}$$

with x being the n dimensional state vector, t is the time, ϕ is a parameter vector, for calculational purposes either one or two dimensional but in general p dimensional. The analysis of such a model is concerned with the solutions of the model. By solution in this context is meant the stationary solution to the model equation A.1:

$$0 = f(x_0,t_0,\phi_0) \tag{A.2}$$

By varying the free parameter ϕ the solution to A.2 changes. Some systems undergo stability changes as the parameter, ϕ, is changes. A stability change occurs when the eigenvalues of the linearized model equations cross the imaginary axis in the complex plane. At a solution to equation A.2 the Jacobian is evaluated

$$J = \frac{df}{dx}|_{(x_0,t_0,\phi_0)} \tag{A.3}$$

associated with the Jacobian are the eigenvalues, λ. As known from linear theory the stability of a process is determined by the real part of eigenvalues of the Jacobian matrix, if the real part of the eigenvalues are all negative the solution is stable and if the real part of at least one eigenvalue is positive the solution is unstable. There are some classical bifurcations which will be described in the next section and indications to where they typically are encountered in chemical engineering problems.

A.3 Bifurcation types

In mathematical terms the simplest bifurcation is the fold bifurcation and it can be described by a simple one dimensional differential equation

$$\dot{x} = \phi - x^2 \tag{A.4}$$

which is defined for $\phi \in [0;\infty[$. The solution to the model equation A.4 is shown in figure A.1

The figure shows what happens as the parameter ϕ is varied around 0. The solution changes from negative to positive and the solution changes stability from stable to unstable. This means that if the system is integrated the solution is determined by the initial condition. If the parameter ϕ is set to 4 then

$$x(t=0) \in [-\infty;2[; \Rightarrow x(t) \to -\infty \tag{A.5}$$

$$x(t=0) = -2; \Rightarrow x(t) = -2 \tag{A.6}$$

$$x(t=0) \in [2;\infty[; \Rightarrow x(t) \to 2 \tag{A.7}$$

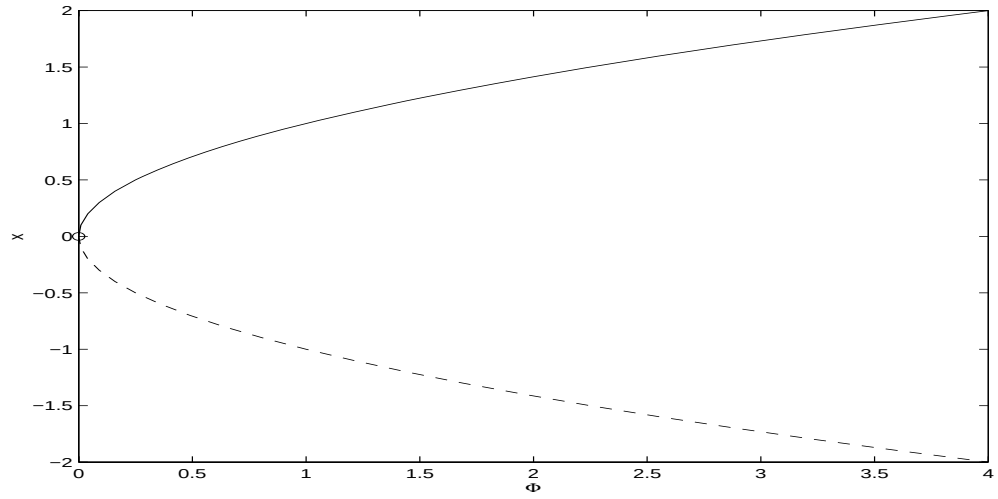

Figure A.1: Fold bifurcation of the model in equation A.4, full line indicates stable steady state, dashed line indicates unstable steady state, the circle is the bifurcation point.

As can be seen from both figure A.1 and the solutions above there exist multiple solutions for equation A.4. This is a specific nonlinear feature, multiple solutions does not exists for linear systems. The point where the transition occurs is called the bifurcation point, in figure A.1 the point $(\phi, x) = (0, 0)$. As the parameter ϕ is changed the eigenvalue of the linearised model equation is changing. At the bifurcation point the eigenvalue is crossing the imaginary axis in the complex plane. This is a characteristic feature of fold bifurcations also of multidimensional systems, where a fold occurs if a single real eigenvalue cross the imaginary axis.

A phenomenon which is related to the occurrence of fold bifurcations is hysteresis which can be explained rather simple in figure A.2. Assume the process is operated with a low value of the parameter ϕ and it is increased. As ϕ passes the point A, the steady state of the process disappears and it will drop down on the low value steady state. To try to recover from this the value of ϕ is decreased but the process keeps running at the low value steady state untill ϕ passes the value at B. Then the low value steady state disappears and the process returns at the high value steady state. This is what is meant by hysteresis, the process does not follow the same route when the parameter is increased as when it is decreased. The process in figure A.2 has two fold bifurcation A,B and the process is unstable on the branch between the two bifurcation points.

The fold bifurcation is also known under different names e.g. limit point and saddle-node bifurcation.

The other classical type of bifurcation is a Hopf bifurcation, which is characterized by the movement of a complex conjugated pair of eigenvalues across the imaginary axis. The model for such a system will necessarily be of at least second order since two eigenvalues are involved in the transition.

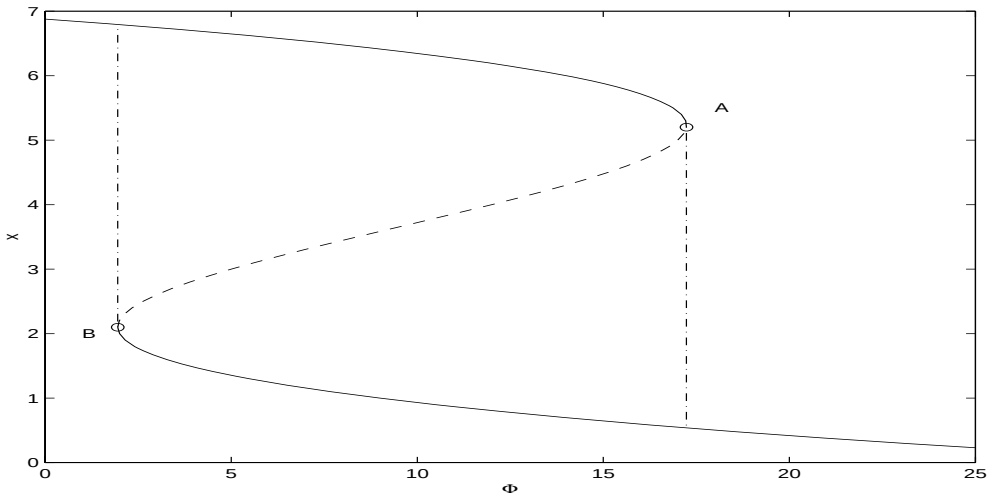

Figure A.2: Bifurcation curve, hysteresis, full line indicates stable steady state, dashed line indicates unstable steady state, the circle is the bifurcation point.

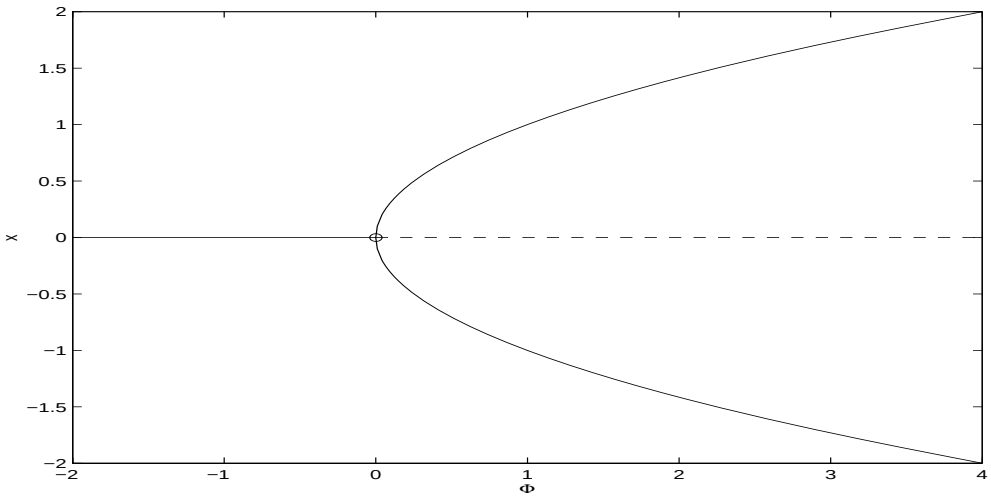

Figure A.3: Hopf bifurcation, state space plot, full line indicates stable steady state, dashed line indicates unstable steady state, the circle is the bifurcation point.

At the bifurcation point the eigenvalues are located at the imaginary axis. On one side of the bifurcation the eigenvalues are stable (negative real part) and one stable steady state exists. On the other side of the bifurcation point the steady state solution have become unstable and a new stable solution appears. This new solution is oscillatory with increasing amplitude as the bifurcation parameter ϕ is increased. This means that sustained oscillations will be observed as a result in time domain.

In chemical engineering these phenomenon have been shown for many different processes. The fold bifurcation is typically seen in tank reactors and Hopf bifurcations are seen in e.g. recycle systems which typically occurs due to integration of process equipment. One resent example is a fixed bed reactor and a heat exchanger which is integrated with recycle of either energy or mass, see Recke (1998). It have been demonstrated that the process can exhibit complicated behaviour.

Another example is sustained oscillations in the glycolysis of microorganisms, as described by Richard et al. (1995) and for some of the early work in the study of oscillations in microorganisms by Golbeter and Caplan (1976). The understanding of the oscillations can be improved by using theory of bifurcation analysis in addition to searching for the explanation in the biochemistry.

A.4 Fold bifurcation - Examples

In many chemical processes there is operational problems due to the existence of hysteresis, that is when a process parameter is varied in one direction one trajectory is followed. If the parameter then is varied in the other direction another trajectory is followed. This means that multiple stable steady states exist in some region of the operating window.

One simple textbook example is an adiabatic CSTR, see Fogler (1992) chapter 8, where the coupling of the energy balance and the mole balance give rise to multiple steady states. Another example is an auto catalytic isothermal reaction in a CSTR, where it is the coupling between the reactor design/operation and the reaction kinetics which give rise to the operational problems. The general characteristics for these examples is the coupling between two process characteristics, one is a bending curve, (R), which is fixed by physics/chemistry and a straight line, (L) which can be moved either parallel or by varying the slope, see figure A.4, typically by changing a design/operational variable. The intersection points in the figure are the steady state solutions of the process equations ($R = L$). The points A and C are the stable steady states and the point B is an unstable solution to the equations. If the straight curve is varied such that it intersects with one of the points D or E then a small change will cause that there is only one steady state solution to the process equations, which is located far from the points D and E. These points are called Fold Bifurcation points. Another way to illustrate how the steady state solution is changing due to changes in the residence time is shown in figure A.5. The points on the steady state curve between D and E are all unstable steady state. This means that if one is

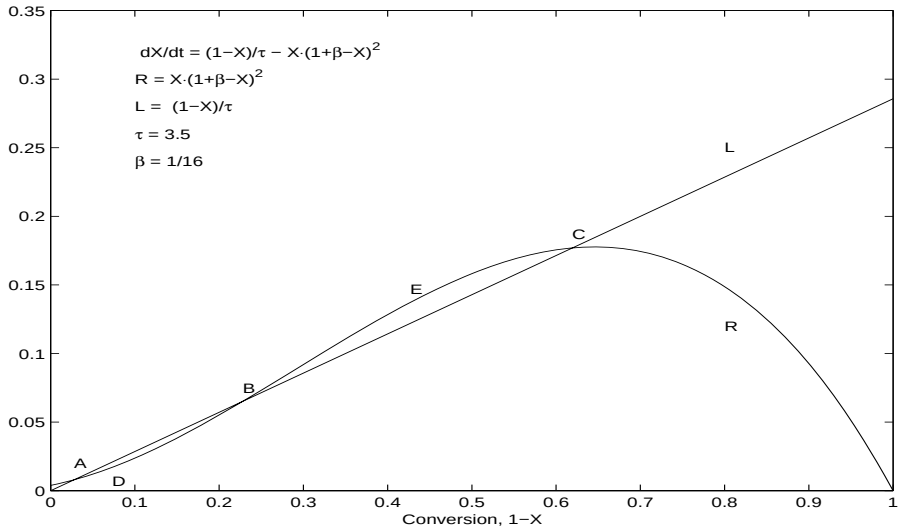

Figure A.4: Coupling between two variables

starting at point B the process will eventually end up at either A or C. In figure A.5 at the point E the steady state solution folds back.

The way to analyse a given system is either by hand or by numerical routines. The analytical approach is only feasible for low dimensional systems. The numerical routines will be described in later parts. Here is given two small examples, one is an adiabatic Continuous Stirred Tank Reactor (CSTR), with a single reaction. The other example is a fermentation system.

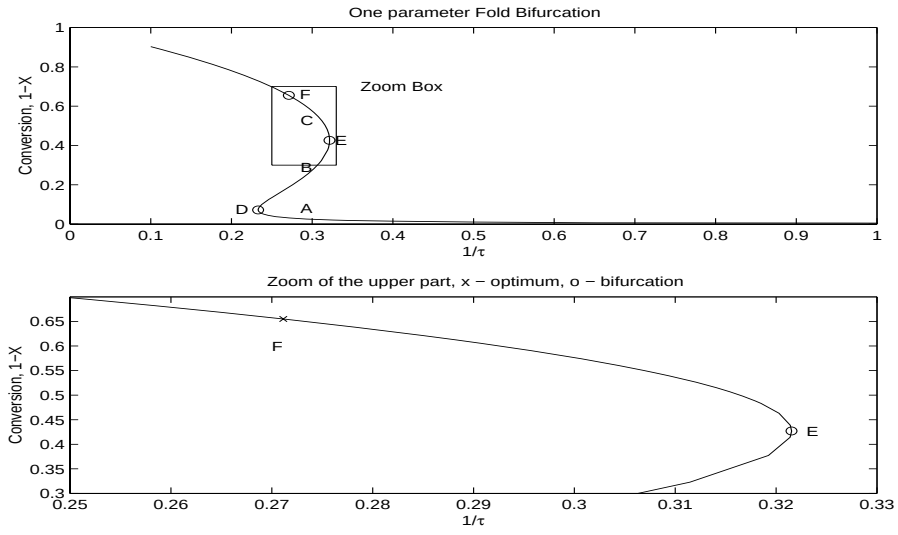

Figure A.5: Bifurcation diagram, steady state solution of the adiabatic CSTR example in section A.4.1

A.4.1 Adiabatic CSTR

The reaction is given by

$$A \rightarrow B \tag{A.8}$$

and the model is given by one differential equation, since there is a simple relation between the concentration of A and B

$$x = \frac{C_A}{C_{A_0}} \quad (1-x) = \frac{C_B}{C_{A_0}} \tag{A.9}$$

The model is given by

$$\dot{x} = \frac{1-x}{\tau} - x(1+\beta-x)^2 \tag{A.10}$$

The parameter $\beta = 1/16$ is a model constant related to the catalyst and τ is the residence time in the reactor, defined by $\tau = V/F$, where V and F is volume and flow rate.

One way to analyse the model is to split it into two parts, a kinetic part and a reactor part. The reactor part is

$$L = \frac{1-x}{\tau} \tag{A.11}$$

and the kinetic part is given by

$$R = x(1+\beta-x)^2 \tag{A.12}$$

These two equations are plotted in the same figure as a function of $1-x$, the conversion of A, this is shown in figure A.4

The plot shows that by varying the τ the slope of the L curve is changed, for low and high values there are only one intersection. The intersection corresponds to steady state, since $\dot{x} = L - R$. For a range of τ where there are three intersections, the steady states denoted A and C in figure A.4 are stable, but the middle one, B is unstable. If the value of τ is varied over a range from zero to above the region where only one steady state exists, the result is as shown in figure A.5.

Another way to find the curve is to solve the equation $L - R = 0$ for $x(\tau)$. This is not straightforward since there are three steady state for some values of τ. Instead it is possible to find $\tau(x)$ which is just as good for plotting purpose.

$$\tau(x) = \frac{1-x}{x(1+\beta-x)^2} \tag{A.13}$$

Figure A.5 shows a typical result of bifurcation analysis of a CSTR. When the process is to be operated, there are two obvious objectives. One can be to have maximal conversion and the other is to have maximal productivity. The analysis is similar for the two cases, here is the case of maximal productivity which

can be described as $\max(1-x)/\tau$. To simplify the analysis the problem is reformulated by exchanging the residence time, τ with its counter part dilution rate, $D = 1/\tau$. Then the problem is as follows

$$\max_D (1-x)D \quad \text{or} \quad \max_x (1-x)D \tag{A.14}$$

$$s.t. \quad \dot{x} = (1-x)D - x(1+\beta-x)^2 = 0 \tag{A.15}$$

The maximum is found where the following two conditions hold

$$\frac{d}{dx}\left((1-x)D\right) = 0 \tag{A.16}$$

$$\frac{d^2}{dx^2}\left((1-x)D\right) < 0 \tag{A.17}$$

These conditions will ensure that the point will be an extremum, by equation A.16, and equation A.17 will ensure that the point is a maximum. By inserting $D = 1/\tau$ from equation A.13 into equation A.16 one obtain the following

$$\frac{d}{dx}\left((1-x)D\right) = \frac{d}{dx}\left(x(1+\beta-x)^2\right) = (1+\beta-x)^2 - 2x(1+\beta-x) = 0 \tag{A.18}$$

The solution is given by $x = (1+\beta)/3$, when $\beta = 1/16$ then $x = 17/48$. This point corresponds to the point where the reaction rate is at its maximum, this can also be seen from the optimization equation A.18, where the result corresponds to optimizing the reaction rate (second term of equation A.18). As can be seen form figure A.4 the optimal reaction rate corresponds to a point near the point C and therefore the optimal point of operation is in the region where there exists multiple solutions.

A.4.2 Fermentation system - simple case

In this section the conditions that are necessary for the fold bifurcation to appear in a simple fermentation system are established. The system in consideration is an unstructured model for fermentation of a microorganism with one limiting substrate. The model is given by the following equations

$$\dot{s} = -\mu(s)x + (s_f - s)D \tag{A.19}$$

$$\dot{x} = \mu(s)x - xD \tag{A.20}$$

s denotes the substrate and x denotes the biomass concentrations, $\mu(s)$ is the biomass growth rate with respect to the substrate concentration and s_f is the substrate concentration in the feed. The steady state of the process is given by the solution of

$$\mu(s_s) = D \tag{A.21}$$

$$x_s = (s_f - s_s) \tag{A.22}$$

To do the analysis on a multidimensional system like this it is necessary to calculate the Jacobian of the model equations in A.20,

$$J = \begin{bmatrix} -\frac{d\mu(s)}{ds}x - D & -\mu(s) \\ \frac{d\mu(s)}{ds} & \mu(s) - D \end{bmatrix} \tag{A.23}$$

When evaluated at steady state the result is simplified to

$$J = \begin{bmatrix} -\frac{d\mu(s)}{ds}|_{s_s}x_s - D & D \\ \frac{d\mu(s)}{ds}|_{s_s} & 0 \end{bmatrix} \tag{A.24}$$

The condition for having a fold bifurcation is that one of the eigenvalues is located at zero. Since the $(2,2)$ element of the Jacobian is zero the condition reduces to that either one of the elements $(1,2)$ or $(2,1)$ should be zero. The $(1,2)$ element of the Jacobian has only the trivial and non interesting solution $D = 0$. The element in $(2,1)$ $\frac{d\mu(s)}{ds}|_{s_s} = 0$ corresponds to a situation where the growth rate has an extremum, either a minimum or most likely a maximum. One example of a kinetic expression where there is an extremum is growth with substrate inhibition. The classical expression of such a kinetics is

$$\mu(s) = \mu_{max}\frac{s}{K_s + s + K_i s^2} \tag{A.25}$$

The result above can be extended to multisubstrate systems. If one investigate the extension of the model equation for biomass production, which can be written as:

$$\frac{dx}{dt} = \mu(s_1, s_2, \cdots, s_n)x - Dx \tag{A.26}$$

It is assumed that the organism is growing on all substrates, but one is limiting the growth, when calculating the Jacobian the result will be:

$$J_x = \begin{bmatrix} x\frac{d\mu(s_1,s_2,\cdots,s_n)}{ds_1} & x\frac{d\mu(s_1,s_2,\cdots,s_n)}{ds_2} & \cdots & x\frac{d\mu(s_1,s_2,\cdots,s_n)}{ds_n} & 0 \end{bmatrix} \tag{A.27}$$

At the steady state the only substrate that will influence the growth rate is the limiting one (e.g. s_1) and the rest will not affect it since they are nonlimiting. The result of this argument is that the elements in J_x from element 2 to $n+1$ all are zero, implying that the kinetics of the limiting substrate will determine if there is a bifurcation and the criterion is the same as for the more simple case, $x\frac{d\mu(s_1,s_2,\cdots,s_n)}{ds_1} = 0$

This means that if the growth rate is inhibited by the limiting substrate then there is a bifurcation and the operation may be troublesome, depending on the location of the bifurcation on the steady state curve.

A.5 Bifurcation - Numerical aspects

As the models for a process are extended the chance for success in solving the bifurcation problems as presented previously drops drastically. Therefore there

have been developed algorithms which can solve the problems numerically. The problems to solve are either one or two parameter investigations of models as in equation A.1. This description is a procedure for the solution methods which are more or less common in the algorithms that have been used in these investigations. The algorithms are presently developed in academia. At DTU an algorithm have been developed, namely PATH. This program is written in fortran and can perform both one ad two dimensional analysis. An attempt to translate the algorithm into C-code and include DAE model formulation, has been done for the one parameter case, but not for two dimensional analysis. AUTO is a program similar to PATH. Common for all these programs is that they were not able to solve the yeast problem that is described in other parts of this thesis. The reason for this is mainly explained that they did not support solution of stiff differential equations. Another flaw of these algorithms were that the selection of the step length in these algorithms which is related to the condition number of the model, to be able to make small steps around the bifurcation points. This will not be a good choice when the model itself is rather stiff and have a high condition number. The rewrite of the PATH algorithm did support the solution of stiff equations and could solve the problem, but took a lot of steps.

The prefered algorithm that have been used in the investigations is called CONT Kubicek and Schreiber (1997). It supports both one and two dimensional analysis, and it can handle stiff problems. One nice property of this algorithm is the 10 fold increase of the speed of calculation.

A.5.1 One parameter Continuation

Continuation means that one parameter is changed and the solution of the differential equation is investigated over a range of parameter values. The algorithm is

1. Solve $f(x_i, \phi_i) = 0$
2. Check $\prod Re\left(\lambda(\frac{df}{dx}|_{x_i,\phi_i})\right) < \varepsilon \quad \Rightarrow$ Bifurcation
3. Update the free parameter $\phi_i = \phi_{i-1} + h\Delta_\phi$
4. Update $x_i = x_{i-1} + h\Delta_x$, new estimate for solution.
5. Update Step size h if necessary
6. Check if finished, else goto top.

The solution of the RHS of the differential equation will be solved and a number of conditions can be checked. In some algorithms there is the possibility to include interval bounds on the states. The algorithms will search and locate the possible bifurcation points. One important point is the updating of both parameter and state vectors, this can be done by using the tangent between two consecutive points and this is reasonable since the steps are small.

The next point in the investigation of process models, when the continuation have been carried out, is investigation of what will happen with the bifurcation points as another free parameter is varied over a range, i.e. 2-dimensional continuation.

A.5.2 Two parameter bifurcation analysis

The starting point in the investigation of a two parameter bifurcation analysis is a bifurcation point. They can be located as described in section A.5.1. Now there is the extra continuation condition that the eigenvalue $\prod Re(\lambda(\frac{df}{dx}|_{x_i,\phi_i})) < \varepsilon$. In comparison to the algorithm in the previous section this condition was only checked, here it is included in the solution set. By having this condition the algorithm will track the bifurcation point in the parameter space as the two parameters are varied. The algorithms also check whether the number of eigenvalues with real part equal to zero change, i.e. collision of bifurcating solutions, crossing of two real eigenvalues. These types of solutions are seen in some systems and can result in strange static and dynamical behaviour of the process around such points.

A.5.3 Interpretation of Bifurcation analysis

The analysis carried out so far have been only in one parameter, namely the dilution rate of the CSTR. The result of these analysis were found by analytical calculation by hand. A fold bifurcation were found and as discussed in section A.4.1 the three solutions are not all attractive from a production point of view. Two were stable (the upper and the lower) and one were unstable (the middle one). The obvious question is what will happen if the parameter β is changed.

The analysis is carried out by the numerical algorithms and is shown in figure A.6. The indicated areas in figure A.6 is interpreted as a process operated with parameters in area *A* will have the three steady states and the line will indicate the location of the bifurcation points. The area *B* indicates the region where there is only one steady state of the process.

A.6 Summary

Here the basics for understanding the how the fold bifurcation have been presented. The method for analysing the problems have been presented, by analytical method as for the Adiabatic case and a numerical method was applied to the two parameter investigations. The search for a suitable algorithm for solving the problem have been conducted and the best algorithm has proven to be CONT Kubicek and Schreiber (1997). The algorithm is relatively simple to apply for new problems and the sped of calculation is very convincing.

The common thing shown with these examples is that the fold bifurcation occur close to the maximal growth/formation rate. This will usually cause a

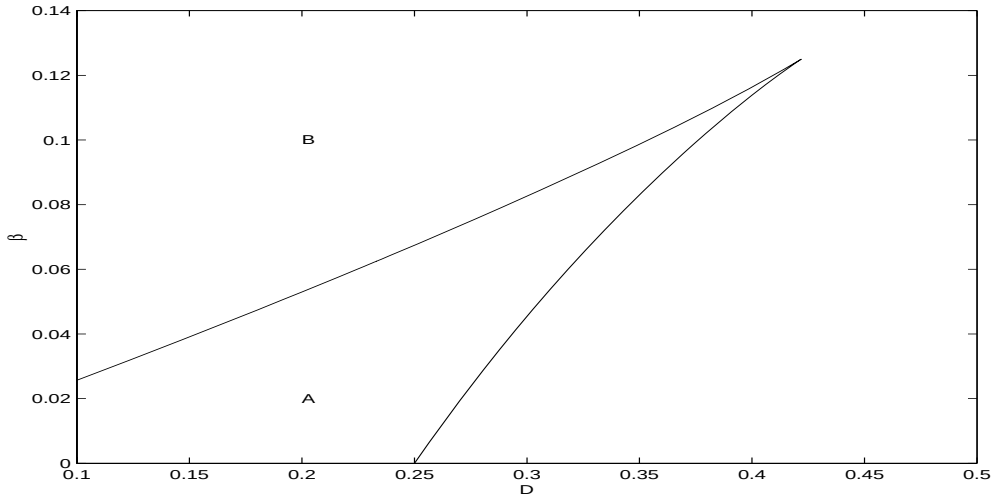

Figure A.6: Two parameter bifurcation analysis of Adiabatic CSTR model. Area A: Three steady states, 2 stable, 1 unstable, Area B: One stable steady state

problem when the process is to be operated in a production facility since the operation in cases where there exists multiple steady states can be complicated.

B

Linearized model equations

Here is given the linearization of the model equations presented in chapter 3 evaluated at the steady state solution determined by a dilution rate of $D = 0.2494871h^{-1}$ and a substrate feed concentration $S_f = 30g/L$. The measurements is the carbon evolution rate and the process input is the dilution rate.

The preliminary investigations of the linearized model at this point is also shown here.

B.1 Linearization results

The model from chapter 3 is linearized here with the input and output selected in chapter 6 for the continuous operating cultivation. The two free variables D and S_f are fixed to $\begin{bmatrix} 0.2494871h^{-1} & 30g/L \end{bmatrix}$ respectively and the corresponding maximal biomass productive steady state is found to be

$$X_{opt} = \begin{bmatrix} 0.0523 \\ 0.0020 \\ 0.0001 \\ 0.0002 \\ 15.1149 \\ 0.4054 \\ 0.0056 \\ 0.0016 \end{bmatrix} \tag{B.1}$$

The model equations is linearized and the steady state in inserted for evaluation. The linear state space formulation will be

$$\dot{x} = Ax + Bu \tag{B.2}$$

$$y = Cx \tag{B.3}$$

Matrices of the linearized process model at the optimal biomass production conditions.

$$A = \begin{bmatrix} -33.8030 & 0 & -47.0787 & 0 & -0.4943 & -18.4314 & 0 & 0 \\ 11.3255 & -691.6545 & 49.0456 & 0 & 0.0000 & 0.0012 & 0 & 0 \\ 0 & 291.6342 & -135.6578 & 9.7489 & 0.0000 & 0.0000 & -105.8850 & -0.5826 \\ 0.0174 & 0 & 42.8789 & -10.4371 & 0.0000 & 0.0001 & -0.1617 & 0.6088 \\ 18.4522 & 0 & 0.1223 & 0 & 0.0000 & 9.3025 & 0.1374 & 0 \\ 0.6787 & 0 & 0.0027 & -0.6058 & 0 & -0.2495 & 0.0054 & 0 \\ -0.0541 & 0 & 0.0000 & 0.0001 & 0 & 0.0000 & -0.2495 & 0 \\ -0.0106 & 0 & 0.0021 & 0 & 0 & 0.0000 & 0.0000 & -0.2495 \end{bmatrix} \tag{B.4}$$

$$B = \begin{bmatrix} 29.9477 \\ -0.0020 \\ -0.0001 \\ -0.0002 \\ -15.1149 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{B.5}$$

$$C = \begin{bmatrix} -0.1023 & 30.0259 & 12.1558 & 0 & 0 & 0.6617 & 13.9940 & 0 \end{bmatrix} \tag{B.6}$$

B.2 Stability of linearized models

The stability of the process is determined by the eigenvalues of the A matrix. The eigenvalues of the A matrix is

$$\lambda = \begin{bmatrix} -716.6891 \\ -111.7237 \\ -35.5808 \\ -6.9268 \\ -0.8437 \\ -0.0378 \\ -0.2495 \\ -0.2495 \end{bmatrix} \tag{B.7}$$

As can be seen the eigenvalues are all negative thus the process is stable at this operation point. The linear model equations will result in a transfer function determined by

$$G(s) = C(Is - A)^{-1}B \tag{B.8}$$

The process transmission zeros can be determined for the given point, and is found to be

$$Z = \begin{bmatrix} 2714.4 \\ -276.31 \\ -8.8263 \\ -0.9452 \\ -0.1036 \\ -0.2495 \\ -0.2495 \end{bmatrix} \tag{B.9}$$

From the transmission zeros it is noticed that there is one positive transmission zero which imply that there is an inverse response of the model at this point. The process gain is determined to $K_p = -3.1248$. Based on the transfer function it is possible to calculate the frequency response which is illustrated by the bode plot shown in figure B.1 As indicated in the bode diagram of the open loop transfer function has a gain margin of 45.3 dB. Since the process gain margin is equivalent to the ultimate gain $K_u = 45.3dB = 184$ and $K_u K_p = 575.2$ Ziegler-Nichols ultimate sensitivity method for tuning feedback controllers is not applicable.

If the next point from the continuation which is located on the other side of the fold bifurcation is selected the steady state solution is given by

$$X0 = \begin{bmatrix} 0.0523 \\ 0.0020 \\ 0.0001 \\ 0.0002 \\ 15.1142 \\ 0.4053 \\ 0.0056 \\ 0.0016 \end{bmatrix} \tag{B.10}$$

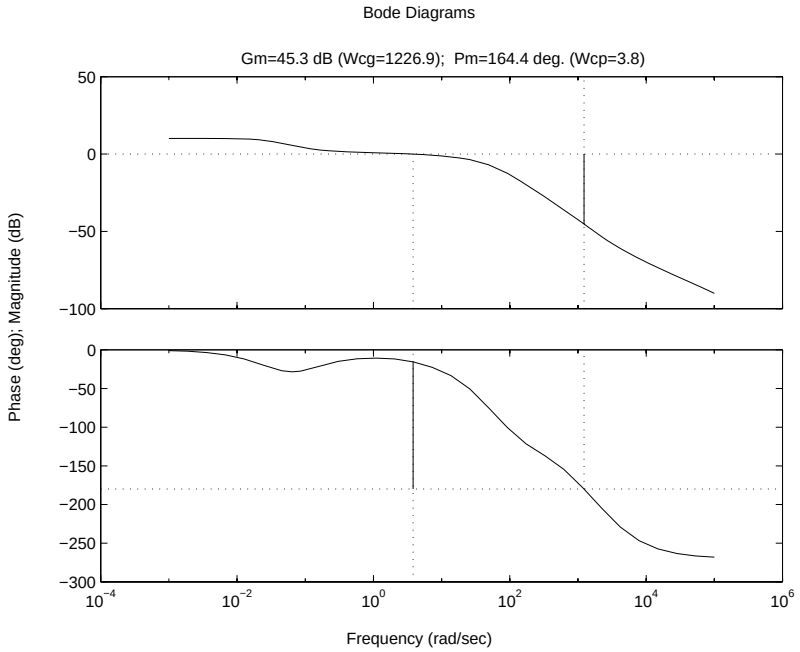

Figure B.1: Bode plot of the open loop transfer function from equation B.8

Then based on this point the matrices A, B and C are calculated and the transfer function is constructed as in equation B.8. The open loop bode plot is shown in figure B.2 it reveals that the point is unstable since the gain margin is negative. This could also have been seen from calculation of the eigenvalues of the corresponding A matrix The closed loop bode plot is shown in figure B.3 The proportional controller is able to stabilize the unstable state and obtain a high gain margin, similar to that of the stable steady state.

B.3 Summary

The linearization of the model equations from chapter 3 is performed for both a stable and an unstable point and the difference bode plots is shown. The unstable point is then stabilized by a proportional feedback controller.

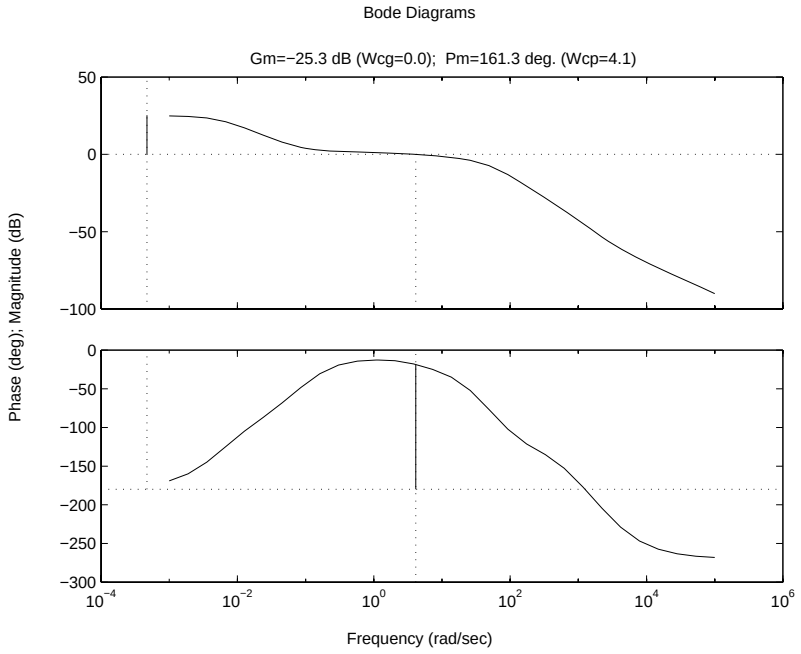

Figure B.2: Bode diagram of the open loop process when the process is evaluated at the unstable steady state

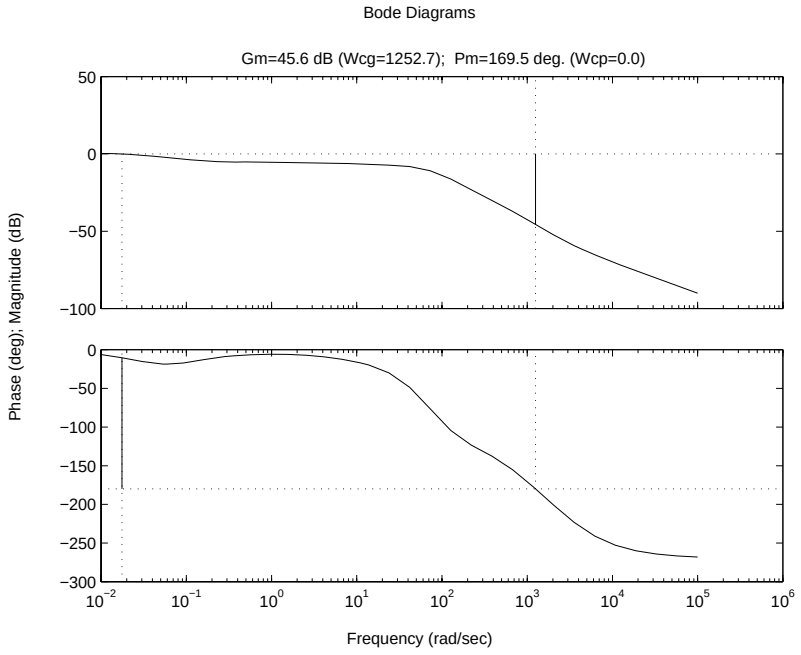

Figure B.3: Bode diagram of the closed loop process when the process is evaluated at the unstable steady state, the process is operated with a proportional controller with a gain of 1

C

A Review on Process Controllability

Kurt Pedersen†, Sten Bay Jørgensen†and Sigurd Skogestad‡
†Department of Chemical Engineering
Technical University of Denmark
Lyngby, Denmark
‡Department of Chemical Engineering
NTNU
Trondheim, Norway

This paper is an unpublished internal report, which is produced within the EU CAPE.NET project (working group 3). Any comments you may have are welcome at: sbj@kt.dtu.dk

Definitions of terms used within process flexibility design and controlability assessment are given and their relations discussed. Methods for controlability evaluation and methodologies for integrating these into the design process are reviewed for some modes of operation. Basically two types of evaluation and design methods prevail; One type is based upon linear model analysis, whereas another type is based upon physical chemical insight and thus provides nonlinear information.
In process design practice the plant piping and instrumentation diagram evolves iteratively using mainly experience and process reasoning to address questions related to plant controlability. It would be desirable to be able to address such questions more quantitatively at different abstraction levels during process design such that controlability evaluation can be integrated into the design process.

C.1 Introduction

From the introduction of the first version of general purpose process controllers during the first quarter of this century, control have developed to become an indispensable part of process operation. In fact, today it would be inconceivable to satisfy environmental requirements and hence to operate most chemical processes without control. This trend develops even further as the plant heat and mass flows become more tightly integrated and as plants are optimised both

in design and operation such that nonlinear features are exploited to a higher extent. The gradual development and incorporation of control methodologies have probably contributed to the present state of ad hoc design habits, where the P&I-diagram at best is developed iteratively during the plant design. In this type of procedure little regard seems to be paid to plant controlability aspects eventhough the reason for control is precisely that of ensuring that the plant is controllable subject to known demand and disturbance variations and also subject to all uncertainties imposed during plant design.
Plant design evolves as a sequence of decisions and evaluations. Traditionally one of the early plant design decisions is to select the pertinent process operation mode, i.e. continuous, periodic or batch. This issue should ideally be settled fairly early during the design procedure as the choice has important ramifications on several steps in the design procedure. The periodic and batch operation modes are dynamic hence their design procedure is somewhat different between them and especially different from the static design procedure used for traditional design of continuous plants. Similar differences are relevant for controlability. Thus the review will also address differences between controlability of these operating modes. Note that start-up and shut-down may be considered as special cases of fed-batch or batch operation. Controllability assessment precedes the control system design during the plant design procedure. Thus the controlability assessment deals with whether the plant is controllable, and ideally also with what is the achievable controlled performance of the plant. How the plant actually is controlled is subsequently dealt with during control structuring and controller design.
The purpose of this paper is firstly to review concepts of controlability and in relation to similar concepts. Secondly methodologies for assessing controlability are reviewed in an attempt to get closer to integrating controlability assessment into plant design and especially into the early design phases and thereby facilitate the total design procedure.

C.1.0.1 Outline

First one of the key issues in defining the complexity of the controlability assessment, i.e. the plant operation mode is discussed. Subsequently some of the basic terms used above will be discussed and some definitions will be provided. Then a brief general review of controlability is given in section 3. Methods for evaluating controlability in linear and in nonlinear plants are presented in section 4. Relations between dynamic flexibility and controlability are discussed in section 5. Methods for integrating controlability evaluation into design procedures are briefly addressed in section 6. An example illustrating controlability measures is also briefly discussed before the conclusions.

C.2 Concepts and Problem definition

To approach the concept of controlability and how it may be assessed a number prerequisites and definitions may be useful. The definitions involve a number of related concepts used to describe specific aspects of plant behaviour and include also a discussion of the impact of process operation modes upon the controlability problem. In addition relations between controlability and plant and control design are discussed to approach how controlability may be incorporated earlier into plant design.

C.2.1 Basic Concepts

Traditionally plant design was an entirely sequential discipline, where the control design was carried out after the plant had been designed. Today plant design is viewed as an iterative procedure where the P&I-diagram also is developed iteratively. Controllability evaluation is intended to be relevant also at earlier design stages than the P&I-diagram is used today.
During plant design a number of basic plant performance requirements have to be ensured in order to obtain a design which provides acceptable operational performance.
Operability: is the ability of the plant to provide acceptable static and dynamic operational performance. Operability includes flexibility, switchability, controlability and several other issues.
Flexibility is the ability to obtain feasible steady state operation at a number of given operating points, i.e. over a range of uncertain conditions. These uncertain conditions can be defined from expected variations in rawmaterial and in process performance.
Switchability is the ability to switch between operating points. The main issues are dynamic feasibility and safety. For service plants fast switching may be desirable to minimize loss of product and energy consumption.
With the above definitions the commonly used term feasibility has both a static aspect, which is incorporated into flexibility and a dynamic aspect which is part of switchability. Since the methods for flexibility evaluation have developed also into dynamic flexibility it seems relevant also to consider their relations to controlability.

C.2.2 Process Operation Modes

The actual operating mode of a process significantly influences the control problem requirements. These differences may be briefly summarized as follows:

- Continuous operation implies control around a constant set of steady state values. Thus a regulation problem results which also should involve flexibility analysis and possibly also switchability analysis.

- Start-up, Shut-down and Semicontinuous operations imply control along a desired trajectory. Often the desired trajectory is of high order or even exponential. Thus a path following servo problem results which may impose a high order load upon the control system. Achieving good control performance may be very difficult using simple control.

- Periodic operation implies control along a periodically repeated trajectory, e.g. a limit cycle imposed by the operating conditions in the case of repeated batch operation. If the trajectory is simple, e.g. purely sinusoidal, then this servo problem imposes a modest load upon the control system, but most often a nonlinear trajectory is required.

The main difference between the regulation and the servo problems is that the regulation problem often may be solved by considering constant loads, whereas the servo problem imposes higher order loads upon the control system. Satisfaction of high order loads may be difficult and require extensive usage of feedforward. During control system design a limit to the allowed order of this load is imposed, this limit of course reflects the designed performance. For all types of the above control problems it is highly desirable to be able to determine the achievable control performance at an early design stage. That is to determine 'How controllable - if at all - is this design?'. Thus it is relevant to attempt to define what is meant by controlability and how it may be determined? Clearly methods for determining controlability depend upon the performance specifications, which also depend upon the operation type, and upon the plant model complexity. To attempt to simplify the discussion some purpose based definitions for controlability are reviewed before commenting upon relations to more mathematical, or control theory based, definitions.

C.3 Controllability

Controllability is used with several different meanings in literature. Below a few of these will be discussed and related to the way controlability may be used during process design.
Two main approaches have been taken to define controlability. One is based upon a goal or purpose oriented view while the other is based upon a mathematical or state space oriented view. Some of the goal oriented definitions are discussed below and complemented with a basic mathematically oriented definition to indicate how a practically useful definition has been reached.
Ziegler and Nichols (1943) in dealing with continuously operated processes defined controlability as:
The ability of the process to achieve and maintain the desired equilibrium value
Rosenbrock (1970) defined controlability in more general terms:
A system is called controllable if it is possible to achieve the specified aims of control, whatever these may be. By extension, the system is said to be more or less controllable according to the ease or difficulty of exerting control.
Thus controlability may be viewed as a property of the plant which indicates

how easy it is to control the plant to achieve the desired performance. Rosenbrock (1970) introduced the term functional controlability:
The system is functionally controllable if given any suitable vector y of output functions defined for $t > 0$, there exists a vector u of inputs defined for $t > 0$ which generates the output vector y from the initial condition $x(0) = 0$
However functional controlability only provides a yes/no type answer, and gives no measure of achievable performance in case the process is not functionally controllable. Dynamic resilience was introduced by Morari (1983) as *the quality of the regulatory and servo behaviour which can be obtained by feedback*. Thus this concept is closely related to functional controlability, but dynamic resilience also include a quality measure of the achievable performance independent of the controller.

In control theory literature the term controlability has only little connection to the ease with which a plant can be controlled. The following definition is termed 'state controlability':
A state is termed controllable if for any initial state $x(0) = x_0$, any time t_1 and final state x_1, there exists an input $u(t)$ such that $x(t_1) = x_1$
In control theory literature a system in termed controllable if all states of the system are 'state controllable'.
Analogously 'state observability' is defined:
A state $x(t)$ of a process is termed observable at some given t if knowledge of the input $u(t')$ and $y(t')$ over a finite time $t_0 < t' < t$ completely determines $x(t)$.
A process is completely observable if all states are state observable. The above definitions of controlability and observability are often referred to as 'Kalman controlability and observability' due to their inventor Kalman (1960).
Kalman also showed that a linear model may be decomposed into its controllable/uncontrollable and observable/unobservable parts using a similarity transformation.

In practice however it may be difficult to obtain acceptable control of all states in a process, even if all states are controllable. However any unstable state must be both state controllable and state observable, in order to close feedback paths around it and thereby stabilize it. Thus the following concepts become useful:
Stabilizability: A process is stabilizable if there exists a controller K which can stabilize all unstable modes. In the linear case this requires that $A - BK$ is stable, i.e. has all its eigenvalues within the left half plane.
Detectability: (Linear case) A process is detectable if the unobservable subspace does not contain any unstable modes. Thus an observer may be constructed for the unstable modes.
Thus it is only necessary to require that the unstable modes are stabilizable and observable. Based upon this requirement **input-output controlability** may be defined (Skogestad (1994) and Skogestad and Postlethwaite (1996)):
Input-output controlability: The ability to achieve acceptable control performance, i.e. to keep the controlled outputs (y) and manipulated inputs (u) within

specified bounds from their setpoints (r), in spite of signal uncertainty (disturbances (d) or noise (n)) and model uncertainty, using available inputs and available measurements.

Input-output controlability is the meaning which is assigned to controlability in the sequel. With this definition controlability is a plant property which reflects how easy it is to control the plant. Plant controlability depends upon many different aspects such as specific plant dynamics, sensitivity to uncertainty, measurement location, actuator constrains and disturbance characteristics. These aspects are discussed below in a brief review of methods for controlability evaluation.

C.4 Controllability evaluation

Controllability may be investigated using three main approaches, a process understanding oriented approach, an optimisation approach and a fundamental nonequlibrium thermodynamics, also called a passivity approach. The first two approaches may be based upon linear models where these apply, whereas the latter approach, which is not yet fully developed, only apply for nonlinear models. These approaches are reviewed below to provide a basis for discussion of approaches for including controlability analysis into process design at an early stage. To partially follow a historical perspective the linearised models are treated first. The focus is on conveying the basic ideas, advantages and limitations of the methods rather than details which may be found in the referenced literature.

C.4.1 Evaluation for linear models

A wealth of controlability measures which are based upon a linear model or a transfer function for the plant have been proposed. These measures may be viewed as analytical or process oriented in that they provide a fundamental understanding of what limits controlability, i.e. the achievable performance of the controlled plant. The process is modelled as: $y(s) = G(s)u(s) + G_d(s)d(s)$, where y is the measured outputs, u the manipulated inputs, d the disturbances while G and G_d are the plant and disturbance transfer functions respectively. The reference inputs are y_{ref}. The control error is e=$y - y_{ref}$. It is assumed that all variables are scaled to be within -1 to 1 by dividing the unscaled signals with their unscaled maximum expected allowed change, Skogestad and Postlethwaite (1996). The conventional feedback controlled process with controller transfer function K is shown in Figure C.1 Plant characteristics which limit achievable performance of feedback control may be outlined based upon the concept of *internal model control* (IMC) Holt and Morari (1985) and Morari and Zafiriou (1989). In IMC a model is used in parallel with the plant, as shown in Figure C.2, such that the feedback signal is the difference between the two outputs. Thus the feedback signal only contains model plant differ-

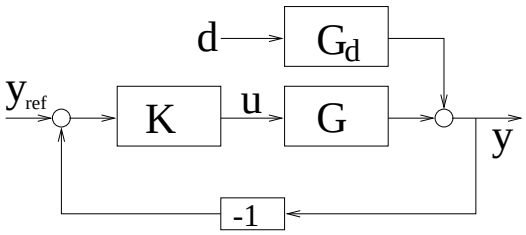

Figure C.1: Conventional feedback blockdiagram for a process with conventional controller (K)

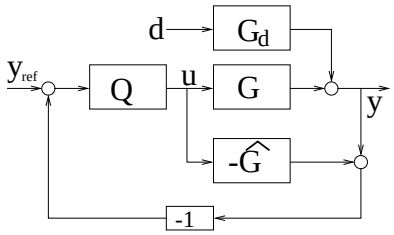

Figure C.2: Internal Model Control feedback blockdiagram with internal model controller (Q)

ences, unmeasured disturbances and noise. Hence if a perfect plant model and perfect knowledge of disturbances are available then no feedback is necessary, i.e. control can be based entirely on feedforward. The transfer function from setpoint to output is $y = GQy_{ref}$, therefore nominal stability is only guaranteed if both the plant and the controller are stable. The intuitively appealing design of an IMC regulator becomes obvious in that perfect control is achieved for $Q = G^{-1}$. Hence any limitation on constructing the plant inverse is a cause for imperfect control, and therefore constitutes a limitation upon achievable control performance. Several limiting phenomena may be listed.

1. Time delays, which when attempted inverted become predictive, which cannot be perfectly accomplished. For a single loop with a time delay of τ the upper bound of the closed loop bandwidth is $\omega_B < \tau^{-1}$. In the multivariable case a lower bound for the settling time for output i is $\tau_i = \min_j p_{ij}$ where p_{ij} is the time delay in the transfer function element g_{ij} (Holt and Morari (1985)). The above authors also propose a lower bound on controlability through the minimum necessary closed loop time delays that are achievable and allow for decoupled control of all outputs: $\rho_j = \max_i(\max(0, \hat{q_{ij}} - \hat{p_{ij}}))$, where $\hat{q_{ij}}$ is the minimum delay in the denominator and $\hat{p_{ij}}$ is the minimum delay in the numerator of element ij of G^{-1}. In some cases an increase of the plant off diagonal time delays may actually improve control. The above result applies for two way decoupling, however if there are large differences between the importance of outputs, one-way decoupling may be justified.

2. Plant zeros in the right half of the complex plane, which become unstable poles in the plant inverse and therefore render the ideal controller unstable. The limiting achievable performance may be illustrated by the upper bound upon the bandwidth for a SISO controller which is approximately $\omega_B < z/2$ where z is a real RHP zero (Skogestad and Postlewaithe(1996)). In the multivariable case it may be possible to move the effect of a RHP zero to a specific output (Morari and Zafiriou 1989). For MIMO systems the frequency range where the high gain can be larger than one is mainly limited to $\omega < |z|$. In MIMO systems the control performance can be poor if a disturbance is aligned with the zero direction. The zero direction is the output singular vector, y_z corresponding to the zero singular value at $s = z$. Thus a requirement for acceptable disturbance rejection is $|y_z^H g_d| < 1$ (Morari et al. (1987) and Skogestad and Postlewaithe(1996)).

3. Unstable poles requires control. To apply IMC design special precautions must be taken (Morari and Zafiriou (1989)). Skogestad and Postlethwaite (1996) argue that for a real RHP pole at p one must expect a closed loop crossover frequency at $\omega_B = 2p$ to achieve acceptable control. Simultaneous presence of both RHP pole and zero requires that the zero is further away from the imaginary axis than the pole to allow for design of a stable controller.

4. Pole excess of the plant transfer function will render the perfect controller improper. A proper controller is obtained by also using a suitable order low pass filter in the controller.

5. Contraints in the manipulated variables. The magnitudes of the disturbances which can be rejected are limited by the actual constraints on the manipulated variables. The static issue here is that of flexibility, but the general problem depends upon the disturbance frequencies. As the controller gain in conventional control approaches infinity to achieve perfect control the control error $e = y - y_r \to 0$ thus $u = G^{-1}y_r + G^{-1}G_d$. Hence if $|[G^{-1}G_d]_{ij}| > 1$, then disturbance j can cause imperfect control of variable i. A plot of the frequency function of the elements of $[G^{-1}G_d]$ provides insight into the possibility of violating input constraints. If constraints are also imposed upon the speed of the actuators that will limit the frequency range within which perfect control performance can be achieved. The most important frequency range for this effect is around the crossover frequency.

6. Model uncertainty will also limit achievable performance. In multivariable plants inversion of the plant transfer matrix becomes difficult the closer to singularity this matrix is. Hence the plant inverse should not be used directly for control. Singularity may be evaluated using a singular value decomposition: $G(i\omega) = U(i\omega)\Sigma(i\omega)V^H(i\omega)$, where the singular values are contained in the diagonal singular value matrix: $\Sigma(i\omega)$. The

condition number is the ratio of the largest and smallest singular values: $\gamma(G) = \frac{\bar{\sigma}(G)}{\underline{\sigma}(G)}$. However this condition number depends upon the scaling. To avoid ambiguity a minimized condition number γ^* is used, where G is pre- and postmultiplied by real diagonal scaling matrices: $\gamma^* = \min_{D_1,D_2} \gamma(D_1 G D_2)$. A large minimized condition number indicates a an ill-conditioned plant.

Another interaction measure is the relative gain array (RGA): $\Lambda(G(s)) = G(s) \times G(s)^{-1}$ where $\times$ denotes element by element multiplication. The ij'th element of Λ can be shown to be the ratio of the open loop gain from input j to output i when all other loops also are open as well, to the gain from input j to output i when all other loops are perfectly controlled. The RGA is very useful as it is scale independent. A relation between the minimized condition number and RGA have been established by Nett and Manousiouthakis (1987): $2\max(||\Lambda(G(i\omega))||_1, ||G(i\omega)||_{\text{inf}}) \leq \gamma^*(G(i\omega)) + \frac{1}{\gamma^*(G(i\omega))}$.

Yu and Luyben (1987) proved that if a single element of G is perturbed from g_{ij} to $g_{P_{ij}} = g_{ij}(1 - 1/\lambda_{ij}$ then the perturbed matrix G_P becomes singular. Thus if an individual element in the plant transfer function has an uncertainty larger than $|1/\lambda_{ij}|$ then the plant may have RHP zeros at the frequency where this occurs. Thus for large relative gain elements extreme sensitivity to uncertainty results. This result has implications for both identification and control, in that plants with large RGA elements will be difficult to identify and also difficult to control due to large sensitivity to uncertainty.

The structured singular value may be used to evaluate how much uncertainty that can be tolerated in the control design. The structured singular value may also be used as an interaction measure to guarantee stability of an overall block diagonal system (Grosidier and Morari (1986)). However this structured singular interaction measure is conservative.

Thus there are a number of controlability measures available which are based upon linear models. Each measure treats one of the control performance limitations and provides information on the qualitative performance limitations, but the measures do not relate directly to the performance requirements.

Closed loop measures are not included since these are not considered in the early design phases. Later during design they may however be used to provide information on controlability limitations and on possible design changes to avoid such limitations.

The main limitations of the linear analysis techniques are that they are based upon input-output models and it may be difficult to relate control performance limitations directly to design variables. Another limitations is the evaluation of influence of uncertainties which is directly related to the specific uncertainty description. Most of the linear analysis tools are based upon frequency domain specifications, whereas often time domain performance specifications are desirable. Finally the application of linear controlability methods require considerable experience as each indicator usually only considers one of the control

performance limitations. Thus there still is a need to further develop even the linear analysis tools and also develop methodologies which further their usage within a process design context.

C.4.2 Optimization methods

Controllability may be evaluated by using an optimization formulation. If the model is assumed linear then many of the above measures may be imposed as constraints to the optimization problem. These methods will be reviewed under nonlinear techniques, due to the possibility for simple extension to nonlinear models, where however convergence proofs are limited.

C.5 Controllability for nonlinear plants

The controlability definition adopted in this paper is based upon the goal or purpose of control, hence this definition applies equally well for nonlinear plants. Nonlinear plants may however exhibit much more complicated dynamic behaviour than linear plants, thus methods for evaluation of controlability may be somewhat different. Even though most plants are truly nonlinear, linear approximations may describe plant behaviour within parts of the operating window. A key property which exemplify nonlinear behaviours is the occurrence of a characteristic change of behaviour within the operating window. Such a change can be the occurrence of multiple steady states for some range of operating parameters, where each of the steady states will have their stability properties. Another characteristic change of behaviour occurs when a previously stable steady state turns unstable. Such behaviours are by now well described for both reactors and separation processes. The operating parameter values at which the characteristic changes in behaviour occur are called bifurcation points. When considering controlability it is of course important to know about such points. One may consider operating the plant outside the region with complicated behaviours. It appears however as if attempting optimal design and operation exploits nonlinear behaviours of process plants (Jørgensen and Jørgensen 1998) such that bifurcations occur somewhere around the optimal operating point. However overdesign may reduce the tendency to complex behaviours (Seader et al. 1990), but clearly an integrated approach to plant design and operation optimization could be most competitive. In such an integrated approach a first step would be to perform a bifurcation analysis to ensure whether and if relevant the types and locations of possible bifurcations. Subsequently each type of behaviour can be further analysed also using linear methods. Thus the controlability analysis is treated through a number of local analysis.

Methods for controlability analysis of nonlinear plants are far less developed than for linear plants. Some aspects for analysing controlability of nonlinear plants are given below, where two types of nonlinear controlability measures

are described which provide some insight, and finally optimization methods are described.

C.5.1 Analytical methods

In nonlinear dynamics a nonlinear inverse may be evaluated. However there are no direct method for quantifying the effect of inverse dynamics. Instead the nonlinear inverse must be analysed. One question which may be directly addressed is whether the inverse of a dynamic system is stable. One approach to this question is to analyze the inverse dynamics, which are usually called *zero dynamics*. The zero dynamics is given by the dynamics of a minimal order realization of the system inverse. The analysis may be grouped into two cases:

1. For constant setpoint the stability of the closed loop with the (right) inverse employed as a controller is completely determined by he stability of the unforced zero dynamics. Daoutidis and Kravaris (1991) describe a nonlinear system as minimum phase if its (unforced) zero dynamics is asymptotically stable, and nonminimum phase if it is unstable.

2. For a reference trajectory tracing the inverse dynamics is driven by the desired system output trajectory and its first $r-1$ derivatives. Where r is the relative degree, which characterises the lowest order derivative of the output y that is explicitly dependent upon u. In this latter case the forced zero dynamics must be evaluated to determine internal stability.

Similarly to the usage of the relative gain array (RGA) for linear systems, the static RGA may be used as a measure of the effect of uncertainty upon controlability. Mijares et al.(1985). The block relative gain (BRG) is extended to nonlinear systems by Manousiouthakis and Nikolaou (1989) to provide a static NBRG and a dynamic version (DNBRG). The static NBRG is shown to be a lower bound for the condition number for the nonlinear system. However it is not clear how these two measure relate to achievable control performance.

C.5.2 A Passivity based Methodology

An interesting new development is passivity based control (Frashman, Viswanath and Ydstie 1998). The merits of this promising methodology are that it is possible to synthesise a guaranteed stabilizing control configuration based upon model information only, where each inventory is controlled. The underlying control design can be relatively simple, such as multiple proportional and integral regulators (inter-)connected with a number of additional measurements to ensure feedforward knowledge about the loads upon the different loops. Two disadvantages are that this far only relatively few processes have been cast into this framework and that the methodology at present requires measurements or estimates of all states, which is the case for all nonlinear control techniques

. A key property of passive systems is that subject to some smoothness requirements a system constructed by an arbitrary interconnection of passive systems is itself a passive system. Controllability aspects of this methodology are straightforward in that presently it is a matter of establishing passivity of the system at hand. Thus there is significant interest in finding representations that render a system passive.

C.5.3 Optimization methods

These methods constitute perhaps the most successful methods in attempting to integrate design and controlability. These methods are reviewed by Perkins and Walsh (1996) and Walsh and Perkins (1996). Here some of the main developments are emphasised to illustrate assessment of controlability. The determination of flexibility is also considered since that methodology has affected assessment of dynamic flexibility, which is related to controlability.

Narraway, Perkins and Barton (1991) provided a measure of the best achievable economic performance as the amount that the operating point must be backed from the optimal operating point to ensure that none of the operating constraints are violated, due to disturbances, thereby affecting controlability. Walsh and Perkins (1992) provided an optimistic bound on disturbance rejection performance by assessing performance of an idealised controller under worst case conditions. The plant performance was limited by both delays and uncertainty. White et al. (1994) evaluated switchability of a proposed design, provided the control system is given. Mohideen et al (1996) extend this to consider operability analysis, control structure selection and controller tuning. Vu, Bahri and Romagnoli (1997) incorporates also operability into the switchability problem. Soroush and Kravaris (1993 a and b) addressed flexibility of operation of batch reactors. they defined flexibility qualitatively as the ability of a reactor to operate according to a predetermined optimal trajectory in the presence of uncertainty. Kuhlmann and Bogle (1998) presents an approach for design of robust controllers which is integrated into a fed-batch design problem formulation. Chenery and Walsh (1998) proposes a linear performance measure to determine a measure of controlability:

$$minJ^F(d) = \min_{K,u_0} \quad \text{s.t.} c(K,u_0,p) \leq 0 \forall p \in P$$

Thus a controller K, a reference operating point u_0 are selected to minimize the objective J while ensuring feasibility for all disturbances p within a bounded set P and satisfying the constraints c. The linear problem is formulated with a linear objective, linear model, linear constraints and LTI controller from the set of stabilizing controllers. Feasibility aspects of the approach are demonstrated on an industrial case study.

C.5.4 Static Flexibility

Design of static processes may be described by the problem:

$$c(d,z,p) \leq 0, \quad \text{where} \quad p_- \leq p \leq p_+ \quad \text{and} \quad z_- \leq z \leq z_+$$

where c represents the process equality and inequality constraints, which are indexed with i. d is the vector of design variables, z the vector of control actuator variables and p the vector of uncertain parameters. Halemane and Grossman (1983) showed that the feasibility problem for processes with the above description is equivalent to the following optimization problem:

$$J^F(d) = \max_p \min_z \max_i f_i(d,z,p)$$

Where $f_i(d,z,p)$ is the description of the process and operational constraints. If $J^F(d) \leq 0$ then the design is feasible. If however $J^F(d) > 0$ then the solution will provide a critical point p^c where the largest violation of the constraints occurs. Swaney and Grossman (1985) extended this work and proposed a flexibility index on top of the feasibility problem with $p^N - \delta\Delta p_- \leq p \leq p^N - \delta\Delta p_+$. The flexibility index $F = \max\delta$ quantifies the ability of the process to operate at other than the nominal operating point. They also showed that under certain convexity assumptions critical points that limit feasibility or flexibility lie on the vertices of the uncertainty space. Saboo et al (1985) also formulated optimization problems to determine static feasibility and flexibility. Grossman and Floudas (1987) exploited the fact that sets of active constraints are limiting design flexibility in their mixed integer linear/nonlinear programming problem. Pistikopoulos and Mazzuchi (1990) introduced a stochastic flexibility index for processes with stochastic parameters. Grossmann and Straub (1991) pointed out that the above two step procedure establishes the flexibility analysis problem:

1. *The feasibility problem:* Determines if a given design can feasible operate over the considered range of uncertainty

2. *The flexibility index problem:* Evaluates a measure to quantify the ability to operate in the presence of uncertainty. Above this measure is establishing the maximum parameter range over which the design can operate feasible.

C.5.5 Dynamic Flexibility

Grossman and Morari (1983) pointed out that several dynamic situations required consideration of process dynamics in flexibility analysis. Dimitriadis and Pistikopoulos (1995) extended the approach of Swaney and Grossman (1985) to dynamic flexibility, where they include time-varying uncertain parameters.

C.5.6 Discussion on Dynamic Flexibility versus Controllability

The static flexibility problem mainly considers the feasibility issue and an index for flexibility. These aspects also are the outset for the dynamic flexibility problem of Dimitriadis and Pistikopoulos (1995), where the key point is to be able to account for parameter variations during operation. Given a solution to the dynamic flexibility issues, then the more strict requirements of achievable control performance may be addressed. Thus in general the set of plants which are statically flexible, includes the set of plants which are dynamically flexible. This set again includes plants which are switchable between a number of operating points, and that set further include plants which are input-output controllable around a given operating point.

Integration of controllability into process design

Several methods for integrating controllability evaluation methods during the different stages of process design have been proposed. One methodology has been to use open loop indicators following the ideas of Morari (1983), and Hovd and Skogestad (1996). This approach was taken, e.g. by Luyben and Floudas (1994), which uses these measures inside a multiobjective optimization formulation.

A simpler approach has been taken, e.g. by Weitz and Lewin (1996) attempts to develop a linear dynamic model from static flowsheet simulation information and simple assumptions concerning process dynamics. This approach is extended in Lewin and Gani (1997).

The methods for controllability evaluation are briefly summarised, before a four step process deisign procedure is given and introduction of controllability measures at each step is briefly discussed.

1. Linear model analysis based methods. These methodologies provide a sound basis for understanding the control problems.(Hovd and Skogestad 1996) and (Larsson and Skogestad 1998).

2. Optimization based methods. A key issue is to propose a superstructure for the problem which is sufficiently rich to ensure that the truely optimal soltion,

in fact is included.

3. Thermodynamically nonlinear model based methods. Passivity based control evaluation

   may be developed to be used at a relatively early design stage to evaluate

   whether a candidate

   flowsheet can achieve a required performance from available measurements

Integration of controllability evaluation methods into plant design is limited by the complexity of the iterative plant design procedure. Some key issues may be illustrated through the following four step design procedure:

1. Conceptual qualitative plant design, where the plant is divided into key stages.

   Based upon reaction routes and thermodynamic separation constraints.

   Design targets are set according to product and waste quality

   requirements.

2. Separation sequences and mass separation agents are selected for each process stage.

   which subsequently can be locally optimised.

3. Integration of stages creates a set of solutions with different

   combinations of process stages. At this level both mass and heat integration are

   relevant and overall design may be optimised.

4. Selection of candidate flowsheet for detailed engineering.

The steps selected in this procedure are illustrative of some of the many decisions which have to be made during process design. Most often each step is carried out several times during a design. The above simplified procedure may be viewed as representing the evolution from more traditional design procedures, where only little optimization and no processintegration might have been achieved during the often very iterative design to more research level design procedures where a design assistant tool may support the designer in making the many design decisions in a consistent manner and where optimiations also guide

the designer to reduce the need for iterations.

To include controllability considerations into process design require the following types
of controllability evaluation tools:

1. Qualitative controllability preferably both at at a total plant level and at each stage.

2. Quantitative controllability evaluation at each stage.

3. Plant wide controllability evaluation.

At the two latter levels 2 and 3 the tools reviewed above
may be used, whereas tools for qualitative controllability evaluation could be
e.g. Hopkins et al. (1998). Linear model based controllability analysis methods
may be incorporated into optimization based process design steps. To include
linear model based analysis tools into qualitative step might be achieved
through combining thermodynamically based qualitative models with qualitative
information from linear qualitative models. However whenever linear analysis is
used this analysis should be preceded by a nonlinear analysis to reveal whether
there are possiblities for multiple solutions, and if so then which solutions
may be stabilised. Thus in those cases nonlinear analysis would be required.

C.6 Example

Application of the above methods for controlability analysis have been exemplified on a number of examples. One study is the energy integrated distillation column at DTU suggested as a case study by Koggersbøl and Jørgensen (1995) where both a simulation and the plant are available for further studies and experiments. The linear methods including optimization based methodologies are exemplified in Hansen et al. (1998). On this plant control configuration is not simple due to the heat pump. The optimization based suggested control configurations are analysed using the linear methods and compared to a heuristically developed control configuration. The optimization based methodology provide control structures which are demonstrated to be better than the heuristical structure. However for the present case it will be most relevant to enable evaluation of controlability at the early stage of design. At present this case study is not elaborated as a design example.

C.7 Conclusions

Controllability is viewed as a plant property which reflects how easy it is to control the plant. Thus plant controlability depends upon many different aspects such as plant design and specific plant dynamics, sensitivity to uncertainty, actuator constraints and disturbance characteristics. These aspects are discussed for different plant models and operating characteristics: In particular it is noted that:

1. Several insightful measures exist for linear plant models. The measures may be directly included in software. However these measures are difficult combine and to use in a meaningful way on different cases.

2. Nonlinear processes require specific case by case controlability analysis.

3. Inventory control combined with passivity based design provides a promising tool not only for controlability analysis, but also for the possibility of integrating design and controlability.

4. Optimization based methods provide also promising tools which with increasing computer power can provide answers on larger and larger problems. The main limitations on these methods are the requirement of a formulated superstructure. Such a structure may easily be limited by the experience of the designer. Optimization provides a possible solution to large problems, but especially for nonconvex problems complexity remains a limitation.

Thus the overall conclusion is that research is needed on several aspects of controlability. These include:

1. How to combine all the linear measures for controlability into a single controlability measure?

2. How to a analyze nonlinear controlability and how to integrate design and controlability evaluation, i.e. to avoid complexity.

3. How to establish passivity of a process plant?

References

M. Akesson, P. Hagander, and J.P. Axelsson. A pulse technique for control of fed-batch fermentations. In *Proceedings of the 1997 IEEE International Conference on Control Applications*, pages 139–44. IEEE, 1997.

M. A. Alexander and T.W. Jefferies. Resperitory effiency and metabolite partioning as regulatory phenomena in yeasts. *Enzyme Microbi. Technol.*, 12: 2–19, Janyary 1990.

M.Y. Andersen, N.H. Pedersen, H. Brabrand, L. Hallager, and S.B. Jørgensen. Regulation of continuous bioreactor near the critical dilution rate using a productostat. *Journal of Biotechnology*, 54:1–14, 1997.

J.P. Axelsson, T. Münch, and B. Sonnleitner. Multiple steady state in continuous cultivations of yeast. In M.N Karim, editor, *Modelling and control of Biotechnical processes*, volume 10, Keystone, Colorado, March 29-April 2 1992. IFAC.

James E Bailey. Mathematical modelling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnology Progress*, 14(1):8–20, 1998.

L.T. Biegler. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Computers and chemical engineering*, 8:243–248, 1984.

J.L. Casti. *Reality ruels: I. Picturing the world in mathematics, the fundamentals*. John Wiley & Sons, 1992a.

J.L. Casti. *Reality ruels: II. Picturing the world in mathematics, the frontier*. John Wiley & Sons, 1992b.

D. Clarke. *Advances in Model-based predictive control*. Oxford science publications, 1994.

J.E. Cuthrell and L.T. Biegler. On the optimization of differential-algebraic process systems. *AIChE Journal*, 33(8):1257 1270, August 1987.

D. Dibiasio and W.A. Weigand. An experimantal investigation of stability and multiplicity of steady states in a biological reactor. *AIChE Journal*, 27(2): 284–292, 1981.

M Faxø. Control of continuous fermentation of *sacchromyces cervisiae*. Master's thesis, Technical University of Denmark, department of Chemical Engineering, 1995.

H.S Fogler. *Elements of Chemical Reaction Engineering*. Prentice Hall International Editions, second edition, 1992.

Bernard Friedland. *Control System Design*. Mcgraw-Hill, 1987.

C.E. Garcia, D. M. Prett, and M. Morari. Model predictive control: Theory and pratice - a survey. *Automatica*, 25(3):335–348, 1989.

A. Golbeter and S.R. Caplan. Oscillatory enzymes. *Annu. Rev. Biochem.*, 5: 449–476, 1976.

J. Guckenheim and P Holmes. *Nonlinear oscilations, dynamical systems, and bifurcationsof vectorfields*. Number 42 in Applied Mathematical Science. Springer Velag, New York, 1983.

V. Hatzimanikatis, M. Emmererling, U. Sauer, and Bailey J.E. Application of methematical tools for metabolic design of microbial ethanol production. *Biotech. Bioeng.*, 58(2 & 3):154–161, 1998.

Vassily Hatzimanikatis, Christodoulos A Floudas, and James E Bailey. Analysis and design of metabolic reaction networks via mixed-integer linear optimization. *AIChE Journal - American Institute of Chemical Engineers*, 42 (5):1277–1292, 1996.

S.B. Jørgensen, H.E. Møller, and M.Y. Andersen. Adaptive control of continuous yeast fermentation, near critical dilution rate. In *Second IFAC symposium on modelling and control of biotechical processes, Keystone, Colorado*, pages 1523–1527, 1992.

J. B. Jørgensen and S. B. Jørgensen. Operational implications of optimality. FOCAPO, Colorado, 1998.

H. Kacser and J.A. Burns. The control of flux. *Symp. Soc. Exp. Biol.*, 27: 65–104, 1973.

P. Karp, M. Riley, S. Paley, A. Pellegrini-Toole, and M Krummenacker. Ecocyc: Electronic encyclopedia of e. coli genes and metabolism. *Nucleic Acids Research*, 27(1):55, 1999.

M. Kubicek and I. Schreiber. Computational analysis of nonlinear dynamical systems. *Zeitschrift fur Angewandte Mathematik und Mechanik*, 77(2): S603–4, 1997.

C. Kuhlmann, I.D.L. Bogle, and Z.S. Chalabi. Roubust operation of fed batch fermenters. *Bioprocess Engineering*, 19:53–59, 1998.

F. Lei, M. Rotbø l, and S.B. Jørgensen. A biochemical structured model for *sacchromyces cerevisiae*. *Internal report*, February 1999.

Frede Lei. Set point tracking of ethanol in continuous cultivation of *saccharomyces cerevisiae*. Reseach report, Technical University of Denmark, Department of Chemical Engineering, September 2000.

J.W. Lengeler. *Bioreactor Engineering*, chapter 10 Basic Concepts of Microbial Physiology and Metabolic Control. European Federation of Biotechnology, Saltsjöbaden, June 1996a.

J.W. Lengeler. *The Phosphoenolpyruvate-dependent Carbonhydrate: Phosphotransferase System (PTS) and control of Carbon Source Utillization*, chapter 11. R.G. Landes Company, 1996b.

L Ljung. *System Identification - Theory for the user.* Prentice-Hall, Englewood Cliffs, N.Y., 1987.

R Luus. Application of dynamic programming to differential-algebraic process systems. *Computers chem Engng*, 17(4):373–377, 1993.

R. Luus. Optimal control af batch reactors by iterative dynamic programming. *J. of process control*, 4(4):218–226, 1994.

D.Q. Mayne and H. Michalska. Receding horizon control of nonlinear systems. *IEEE Trans. A.C.*, 35(7):814–824, 1990.

H Michalska and D.Q. Mayne. Robust receding horizon control of constrained nonlinear systems. *IEEE Trans. A.C.*, 38(11):1623–1633, 1993.

H Møller. *Modelling, Identification and control of a Continuous fermentation with Sacchromyces cerevisiae*. PhD thesis, Department of Chemical Engineering, Technical University of Denmark, 1993.

J. Nielsen and J. Villadsen. Modelling of microbial kinetics. *Chemical Engineering Science*, 47(17/18):4225–4270, 1992.

Jens Nielsen and John Villadsen. *Bioreaction engineering principles*. Plenum Press, 1994.

T. Nørretranders. *Mærk verden : En beretning om bevidsthed*. Gyldendal, 1991.

L.S. Pontryagin, V.G. Bol'tanskii, R.S. Gamkrelidze, and E.F. Mischenko. *The Mathematical Theory of Optimal Processes*. Pergamon, 1964.

E. Postma, W. A. Sheffers, and J.P. van Dijken. Kinetic of growth and glucose transport in glucose limited chemostat cultures of *sacchromyces cerevisiae*. *Yeast*, 5:159–165, 1989.

J. B Rawlings and K.R. Muske. The stability af constrained recceding horizon control. *IEEE Trans. A.C.*, 38(10):1512–1516, 1993.

Bodil Recke. *Nonlinear Dynamics and Control of Chemical Processes*. PhD thesis, Department of Chemical Engineering, Technical University of Denmark, 1998.

Peter Richard, Barbara M. Bakker, Bas Teusink, Karel van Dam, and Hans V. Westerhoff. Acetaldehyde mediates the syncronization of sustained glycolytic oscillations in populations of yeast cells. *European journal of biochemistry*, 1995. In publication.

M. Rieger, O. Kappeli, and A. Fiechter. The role of limited respiration in the incomplete oxidation of glucose by sacchromyces crevisiae. *J. Gen. Microbiol.*, 1983.

M. Rotbøll. Udvikling af model for kontinuert gæring. Master's thesis, Technical University of Denmark, department of Chemical Engineering, 1992. In danish.

Morten Rotbøll and Sten Bay Jørgensen. A biochemical structured model for yeast fermentation. In *European Conference on Biotechnology,*, Florence, 1992.

S.K. Scott and P Gray. *Chemical oscillations and instabilities. Non-linear chemical kinetics*. Clarendon Press, Oxford, England, 1990.

B. Sonnleitner and O. Kappeli. Growth of *sacchromyces cerevisiae* is controlled by its limited respiratory capacity formulation and verification of a hypothesis. *Biotechnology and Bioengineering*, 28:927–937, 1986.

D.E. Steinmeyer and M.L. Shuler. Structured model for *sacchromuces cervisiae*. *Chemical Engineering Science*, 44(9):2017–2030, 1989.

A.P.J. Sweere, J. Giesselbach, R. Barendse, R. de Krieger, G. Honder, and K.C.A.M. Luyben. Modelling the dynamic behaviour of *saccharomyces cervisiae* and its application in control experiments. *Applied Microbiology and Botechnology*, 28:116–127, 1988.

J.M.T. Thompson and H.B. Stewart. *Nonlinear Dynamics and Chaos*. John Wiley and sons, 1986.

D. Tieu, W. R. Cluett, and Penlidis. A comparison of collocation methods for solving dynamic optimization problems. *Computers in chemical engineering*, 19(4):375–381, 1995.

P.M. van den Hof and R.J.P. Schrama. An indirect method for transfer function estimation from closed loop data. *Automatica*, 29(6):1523–1527, 1993.

P.M. van den Hof and R.J.P. Schrama. Identification and control - closed-loop issues. *Automatica*, 31(12):1751–1770, 1995.

J. Villadsen and M.L. Michelsen. *Solution of diffeential equation models by polynomial approximation*. Prentience-Hall, Englewood Cliffs, New Jersey, 1978. 3662 modelanalyse af komplicerede kemiske systemer.

H.K. von Meyenburg. *Katabolit-repression under sprossungszyklus von Saccharomyces cerevisiae*. PhD thesis, Eidg. Techn. Hochshule, Zurich, 1969.

R.F. Weaver and Hedrick. *Genetics*. Wm. C. Brown Publishers, second edition, 1992.

Z.-L. Xiu, A.P. Zeng, and W.-D. Deckwer. Mutiplicity and stability analysis of microorganisms in continuous cluture: Effect of metabolic overflow and growth inhibition. *Biotech. Bioeng.*, 57(3):251–261, 1998.

T.C. Zangirolami, C.L. Johansen, J. Nielsen, and S.B. Jørgensen. Simulation of penicilin production in fed-batch cultivations using a morphologically structured model. *Biotechnology and Bioengineering*, 1997.