

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**Otimização Dinâmica em Tempo Real  
Utilizando Modelos Não-Lineares  
Simplificados**

TESE DE DOUTORADO

Ricardo Guilherme Duraiski

**Porto Alegre**

**2009**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**Otimização Dinâmica em Tempo Real  
Utilizando Modelos Não-Lineares  
Simplificados**

Ricardo Guilherme Duraiski

Tese de doutorado apresentada como requisito  
parcial para obtenção do título de Doutor em  
Engenharia

Área de concentração: Controle de Processos

**Orientadores:**

**Prof. Dr. Jorge Otávio Trierweiler**

**Prof. Dr. Argimiro Resende Secchi**

**Porto Alegre**

**2009**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

A Comissão Examinadora, abaixo assinada, aprova a tese de doutorado *Otimização Dinâmica em Tempo Real Utilizando Modelos Não-Lineares Simplificados*, elaborada por Ricardo Guilherme Duraiski, como requisito parcial para obtenção do Grau de Doutor em Engenharia.

Comissão Examinadora:

---

Prof. Dr. Luís Cláudio Oliveira Lopes.

---

Eng. Dr. Mario Cesar Mello Massa de Campos.

---

Dr. Eng. Pedro Rafael Bolognese Fernandes.

## **Agradecimentos**

Agradeço a todas as pessoas e instituições que contribuíram para o desenvolvimento deste trabalho e para o meu desenvolvimento pessoal durante o Doutorado, em especial:

A Petrobras, a FINEP, ao CNPq e a CAPES pelo financiamento do projeto, bem como a UFRGS por proporcionar todos os recursos materiais e o ambiente de trabalho.

Ao Engenheiro Pedro Rafael Fernandes pela disponibilização do modelo que serviu de exemplo nesta tese

Aos professores Jorge Otávio Trierweiler e Argimiro Resende Secchi por transmitirem tanto de seu conhecimento.

Aos colegas e bolsistas do GIMSCOP, que trabalham no desenvolvimento deste projeto.

A todos os membros de minha família, namorada e amigos pelo apoio e compreensão.

A TriSolutions Soluções em Engenharia, pelo suporte e apoio no desenvolvimento deste trabalho.

## Resumo

Mais que controlar os processos industriais, é necessário fazer com que a sua rentabilidade atinja o máximo possível e isto só é alcançado quando, por meio de algum método ou algoritmo, determina-se a operação ótima do processo. No entanto, a determinação de um ótimo nominal não conduz necessariamente com o ótimo operacional do processo, pois distúrbios de processo e uma série de incertezas podem fazer com que o processo não permaneça otimizado durante toda a operação.

O uso de otimizadores em tempo real corrobora para correção e permanente condução do processo ao seu ponto ótimo. Neste estudo serão discutidas as fundamentações teóricas sobre as quais a otimização dinâmica em tempo real está baseada atualmente, bem como será proposto um novo método de otimização englobando algoritmo, modelos, reconciliação de dados e ajuste de parâmetros.

## **Abstract**

Besides to control the industrial processes, it is necessary to take its profitability reaches the maximum values and this can only be reached when applying some method or algorithm to determine the best operation point of the process. However, the nominal determination of the best operation point does not necessarily takes the process to its best production, because process disturbances and a series of uncertainties can occur and the process does not keeps itself optimized during all the operation.

The real-time optimization corroborates for correction and permanent conduction of the process to its best operation point. In this study it will be discussed the theoretical basis, on which the dynamic real-time optimization is based nowadays, as well as a new optimization method will be proposed including algorithm, models, data reconciliation and adjustment of parameters.

# Sumário

<b>1 Introdução .....</b>	<b>1</b>
1.1 Uma visão geral sobre otimizadores em tempo real .....	2
1.2 Motivação.....	3
1.3 Apresentação da Tese.....	4
1.4 Referências Bibliográficas .....	5
<b>2 Revisão Bibliográfica .....</b>	<b>7</b>
2.1 Princípios da Otimização em Tempo Real.....	7
2.2 Histórico.....	11
2.3 Algoritmos de Otimização .....	14
2.4 Algoritmos Aplicados a RTO e DRTO (Dynamic Real-Time Optimization) .....	14
2.5 Modelos para Otimização .....	16
2.6 A Interação Otimizador-Controlador .....	18
2.7 Tratamento de Incertezas .....	19
2.7.1 Incertezas de mercado.....	19
2.7.2 Incertezas de processo.....	20
2.7.3 Incertezas de medida.....	21
2.7.4 Incertezas de modelo.....	22
2.7.5 Métodos para trabalhar com incertezas.....	22
2.8 Referências Bibliográficas .....	25
<b>3 Algoritmo de Otimização GA-GRG .....</b>	<b>27</b>
<b>Apêndice A - Artigo: “GA-GRG Optimization: A New Algorithm For Multi-Objective Problems Based on the Goal Attainment Method” .....</b>	<b>29</b>
<b>4 Conversão de Modelos.....</b>	<b>55</b>
<b>Apêndice B - Artigo: “Method for Conversion of First Principle Models into Simplified Formulations for Application in Dynamic Real-Time Optimization” .....</b>	<b>57</b>
<b>5 Otimização Dinâmica.....</b>	<b>93</b>
5.1 Proposta de um Otimizador Em Tempo Real .....	93
5.1.1 Visão Geral do Otimizador .....	94
5.1.2 A Simulação do Processo.....	95
5.1.3 O Otimizador.....	95
5.1.4 A Função Objetivo .....	96
5.1.5 O Filtro de Kalman Estendido.....	102
5.2 O Modelo do Reator de Polipropileno .....	106
5.3 Conversão do Modelo Fenomenológico .....	108
5.4 Otimização com Modelos Aproximados Simplificados .....	123
5.5 Conclusões .....	127
5.6 Referências Bibliográficas .....	129
<b>6 Conclusões.....</b>	<b>131</b>

6.1 Considerações Gerais .....	131
<b>Anexo I - Relatório: “Modelagem de Polimerização de Polipropileno via Catalisador Ziegler-Natta em reator CSTR slurry” .....</b>	<b>137</b>
<b>Lista de figuras .....</b>	<b>ix</b>
<b>Lista de tabelas .....</b>	<b>xi</b>

## Lista de figuras

<b>Figura 2.1:</b> Lucro do processo de absorção como função da variável manipulada (relação entre as vazões de óleo e fase aquosa), para diferentes custos da fase aquosa representado pelo parâmetro $\alpha$ . (Lucks 2003).....	10
<b>Figura 5.1:</b> Comparação dos ótimos obtidos pelas duas diferentes formas de otimização multi-objetivo para funções não convexas proveniente de termos convexos.....	99
<b>Figura 5.2:</b> Comparação entre os ótimos obtidos pelo método convencional e minmax (caso 1).....	100
<b>Figura 5.3:</b> Comparação entre os ótimos obtidos pelo método convencional e minmax (caso 2).....	101
<b>Figure 5.4:</b> Perturbações realizadas para determinar as constantes cinéticas do novo catalisador. ....	107
<b>Figure 5.5:</b> Perturbações realizadas para realizar a conversão do modelo fenomenológico em simplificado (sentido direto) .....	110
<b>Figure 5.6:</b> Perturbações realizadas para realizar a conversão do modelo fenomenológico em simplificado (sentido reverso).....	110
<b>Figure 5.7:</b> Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada diferentes das utilizadas na identificação. ....	111
<b>Figure 5.8:</b> Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada diferentes das utilizadas na identificação. Visualização individual de cada perturbação .....	111
<b>Figure 5.9:</b> Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada diferentes das utilizadas na identificação. Amplificação do modelo original.....	112
<b>Figure 5.10:</b> Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada iguais às utilizadas na identificação. ....	112
<b>Figure 5.11:</b> Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada iguais às utilizadas na identificação. Visualização individual das perturbações .....	113
<b>Figure 5.12:</b> Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada iguais às utilizadas na identificação. Resposta do modelo original amplificada. ....	114
<b>Figure 5.13:</b> Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Perturbações diferentes das utilizadas na identificação. ....	114
<b>Figure 5.14:</b> Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Perturbações diferentes das utilizadas na identificação. Ampliação do gráfico da resposta .....	116
<b>Figure 5.15:</b> Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Vazão de catalisador A mantido em 0,2 e Vazão de Hidrogênio modificado de 0,95 para 0,2.....	117

<b>Figure 5.16:</b> Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Vazão de catalisador A levado de 0,2 para 0,95 e Vazão de Hidrogênio modificado de 0,20 para 0,00. ....	117
<b>Figure 5.17:</b> Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações presentes na identificação do modelo aproximado. ....	118
<b>Figure 5.18:</b> Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações presentes na identificação do modelo aproximado. ....	119
<b>Figure 5.19:</b> Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem. ....	120
<b>Figure 5.20:</b> Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem. ....	121
<b>Figure 5.21:</b> Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Terceira perturbação do conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem. ....	121
<b>Figure 5.22:</b> Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Quarta perturbação do conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem. ....	122
<b>Figure 5.23:</b> Comparação das respostas do modelo original e aproximado para modelos baseados em momentos de segunda ordem frente a perturbações presentes no conjunto utilizado para a identificação. ....	122
<b>Figure 5.24:</b> Detalhe da comparação das respostas do modelo original e aproximado para modelos baseados em momentos de segunda ordem frente a perturbações presentes no conjunto utilizado para a identificação. ....	123
<b>Figura 5.25:</b> Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10. $wMI = 1$ , $wXcatA = 0.1$ , $wFH2 = 0.1$ .....	125
<b>Figure 5.26:</b> Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10. $wMI = 3.3$ , $wXcatA = 0.1$ , $wFH2 = 0.1$ .....	126
<b>Figure 5.27:</b> Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10. $wMI = 3.3$ , $wXcatA = 0.1$ , $wFH2 = 0.1$ , $wcFH2 = 3.3$ .....	126
<b>Figure 5.28:</b> Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10. $wMI = 3.3$ , $wXcatA = 0.1$ , $wFH2 = 0.1$ , $wcFH2 = 10$ .....	127

## Lista de tabelas

<b>Tabela 5.1:</b> Pontos da malha retangular utilizados para a conversão do modelo.....	108
<b>Tabela 5.2:</b> Parâmetros de configuração do otimizador.....	125



# Capítulo 1

## Introdução

Desde a revolução industrial, os processos produtivos pelos quais a humanidade obtém os seus bens e produtos de consumo vêm experimentando um avanço considerável e cada vez mais acelerado. Diversos são os fatores que levam os processos produtivos a se desenvolverem tão velozmente. O aumento da demanda de produtos impelido pelo crescimento da população mundial, a forte concorrência pelo mercado, a escassez de recursos naturais para a produção de energia e matéria-prima e a conscientização sobre a preservação do meio ambiente obrigaram as indústrias a aprimorar cada vez mais os seus processos produtivos, levando estes a operar cada vez mais próximos ao limite de sua capacidade. Em uma época em que se exige produtos que reúnem conceitos tão antagônicos, como mais qualidade e menor preço, nenhuma indústria pode mais desperdiçar matéria-prima e energia, ou gerar rejeitos industriais sem restrições, sob pena de ser aniquilada seja pela legislação vigente, seja pela concorrência.

Este conjunto de fatores faz com que os processos industriais se tornem mais complexos e integrados mássica e energeticamente, o que torna seus comportamentos menos intuitivos e com maior dificuldade de prever qual a melhor condição operacional, pois as relações que os governam tornam-se cada vez mais complexas.

A otimização em tempo real de processos vem aliar-se a esta nova visão empresarial, a fim de tornar os processos mais versáteis e capazes de suprir as necessidades das unidades fabris para atender um mercado cada vez mais exigente

e com um comportamento tão variante, da forma mais atrativa do ponto de vista econômico. Tal é a importância da otimização que um mesmo processo industrial, variando-se apenas a sua forma de operação, pode se tornar lucrativo ou não.

Neste capítulo são feitas as considerações iniciais para que se possam entender os desenvolvimentos realizados até aqui, os motivos que os inspiraram, e a proposta de desenvolvimentos futuros.

## 1.1 Uma visão geral sobre otimizadores em tempo real

A camada de otimização de um processo industrial surgiu com o avanço e o crescimento das estruturas de controle. Primordialmente, o controle de processos procura manter o processo funcionando através de um controle regulatório, que se preocupa em, praticamente, enviar ações de controle a atuadores simples baseadas nos objetivos (setpoints) especificados para níveis, pressões e temperaturas de processos. Isto não é o suficiente, pois exige dos operadores um trabalho considerável para, baseados nestas variáveis de controle, tomar decisões para manter a qualidade dos produtos e suas especificações.

Eis que surge então o conceito de controle supervísório. O controle supervísório veio a complementar a automação dos processos industriais no sentido de tornar as tomadas de decisão em relação as variáveis controladas mais intuitivas e visíveis ao operador. Com este advento passou-se a controlar diretamente os objetivos finais do controle tais como qualidade e especificação de produtos. O controle supervísório veio então a complementar o controle regulatório tornando-se uma malha de controle superior a esta na ordem hierárquica.

Contudo, uma nova necessidade surgiu no âmbito industrial em termos dos objetivos finais de se controlar um processo. Certamente, é muito bom que se tenham produtos especificados sendo gerados dentro do processo, mas seria este o objetivo final de uma empresa como uma entidade econômica? Na verdade, um dos objetivos finais de qualquer empresa é inegavelmente o lucro, e entre se ter produtos especificados e lucro existe um *gap* conceitual. Nem sempre um produto especificado é sinônimo de lucro e um esforço adicional se torna necessário neste sentido.

Surge então a necessidade de se otimizar um processo industrial. A otimização se preocupa, de forma geral, em aumentar o lucro da unidade produtiva. No entanto, nem sempre a quantificação do lucro é algo palpável, pois uma unidade

industrial encontra-se sempre introduzida em um contexto mais amplo, composto por várias unidades industriais e o lucro global nem sempre é representado pelo lucro local. Os objetivos da otimização podem ser variados, de acordo com o que a programação de produção, a camada superior à otimização, determinar. Exemplos de objetivos comuns de otimização de unidades industriais podem ser, além do lucro, o mínimo consumo de energia, a máxima conversão de um produto ou vários, a máxima seletividade, a máxima ou mínima especificação, entre outros.

Uma das razões de se fazer otimização *online*, ou em tempo real (RTO), vem do fato de que a determinação do ótimo vem de um modelo matemático do processo, que devido as suas imperfeições pode apresentar divergências em relação ao comportamento real do processo em um dado instante ou intervalo de tempo. Por isso, a correção permanente deste modelo é fundamental para que a determinação do ótimo seja eficaz. Conseqüentemente, muito mais do que aplicar bons algoritmos de otimização, a aplicação de RTO necessita de bons algoritmos de reconciliação de dados e estimação de estados e parâmetros de modelo. As constantes perturbações externas a que os processos estão sujeitos e as variações de características dos equipamentos do processo (acúmulo de impurezas, desativação de catalisadores, perdas de eficiências, etc.) são outras razões importantes para a otimização em tempo real.

## 1.2 Motivação

A otimização em tempo real é um tema muito pesquisado, no entanto, embora se tenha uma gama considerável de pesquisas nesta área, pouco se tem de resultados aplicáveis de forma generalizada a processos industriais e as soluções geradas parecem se aplicar a alguns casos específicos (Cheraghi et al. 1996), (Lucks 2003). O ideal no desenvolvimento de uma tecnologia é que ela possa ser facilmente generalizada e aplicada em diferentes situações como ocorre com controladores PID (Proporcional Integral Derivativo), preditivos lineares, e muitas outras. No entanto, quando se rumo para o campo das descrições não-lineares de processo as soluções parecem ser extremamente específicas para cada caso e, conseqüentemente, morosas e caras.

A dependência de uma descrição ampla e não-linear do processo faz com que soluções de otimização em tempo real não possuam uma generalidade suficientemente grande para se estenderem facilmente para diferentes situações. A busca de uma generalidade e uma aproximação da solução do problema de otimização em tempo real a uma forma fechada é o que motiva o desenvolvimento

desta tese. A associação de técnicas de modelagem não-linear com desenvolvimentos realizados neste trabalho contribuirá para a construção de um otimizador de processos em tempo real não-linear, capaz de ser estendido a diferentes processos através de uma metodologia bem definida.

Objetivando não somente a busca pela generalidade, mas também a obtenção de uma solução de otimização de processos industriais rápida o suficiente para ser aplicada em tempo real, mais do que desenvolver um otimizador, esta tese se propõe a abordar com maior ênfase o aspecto dos modelos não-lineares e sua simplificação a fim de converter modelos fenomenológicos em formulações simplificadas de fácil e rápida avaliação para serem aplicados em otimização e controle avançado.

### **1.3 Apresentação da Tese**

Esta tese de doutorado está dividida em cinco capítulos, dos quais os outros quatro após este de introdução versam sobre os seguintes temas.

No Capítulo 2 apresenta-se a revisão bibliográfica sobre os temas de otimização em tempo real e modelagem não-linear. Atualmente existe uma quantidade significativa de publicações e resultados de pesquisas nestas áreas, de forma que se optou por enfatizar as pesquisas significativas e que dão a base teórica a este trabalho, sem discutir profundamente desenvolvimentos realizados em linhas de pensamento diferentes.

No Capítulo 3 é apresentado um novo algoritmo de otimização desenvolvido especialmente para otimizar funções multiobjetivo da forma MINMAX sendo passível de aplicação a funções mono objetivo. Este algoritmo determina uma direção de busca otimizada para o problema e tem como principais vantagens a transformação do problema de otimização em um problema de função objetivo linear e restrições não-lineares bem como a redução das dimensões do problema numérico a ser resolvido.

No Capítulo 4 é detalhada uma nova forma de conversão de modelos fenomenológicos em modelos simplificados de rápida avaliação. Atualmente, um dos maiores obstáculos à aplicação de otimização em tempo real na indústria é a obtenção de um modelo não-linear, de rápida avaliação, capaz de ser executado inúmeras vezes e em poucos segundos pelo otimizador, que represente com a maior fidelidade possível o comportamento do processo nos mais diferentes pontos de

operação, passíveis de serem atingidos e que possam vir a ser um ponto de operação otimizado. A proposição de uma metodologia de geração de modelos não-lineares rápidos que consigam contemplar estes requisitos, além de fornecer uma representação do comportamento dinâmico do processo, foi um dos tópicos estudados nesta tese.

No Capítulo 5 é mostrada a implementação do otimizador dinâmico em tempo real, o qual associa as duas metodologias desenvolvidas nos capítulos anteriores. São mostrados também os resultados da aplicação do otimizador implementado na otimização da transição de pontos de operação de um reator de polimerização, os quais mostram as vantagens, desvantagens e restrições à utilização deste algoritmo, bem como os desenvolvimentos necessários para que o mesmo se torne aplicável.

Por fim, o Capítulo 6 traz as conclusões e considerações finais a respeito do otimizador como um todo, mostrando os resultados obtidos e a necessidade de novas pesquisas e desenvolvimentos que poderão trazer confiabilidade a esta metodologia de otimização dinâmica em tempo real.

## **1.4 Referências Bibliográficas**

Cheraghi SH, Lim HS, and Motavalli S. (1996). Straightness and Flatness Tolerance Evaluation: An Optimization Approach. *Precision Engineering* 18, 30 - 37.

Lucks D. (2003). Real-Time Optimization in Nonlinear Chemical Processes: Need For Global Optimizer. *AIChE Journal* 49, 2980 - 2983.



## Capítulo 2

### Revisão Bibliográfica

Mais que controlar os processos industriais, é necessário fazer com que a sua rentabilidade atinja o máximo possível e isto só é alcançado quando por meio de algum método ou algoritmo determina-se a operação ótima do processo. No entanto, a determinação de um ótimo nominal não conduz necessariamente o processo a sua produção ótima, pois distúrbios de processo e uma série de incertezas podem fazer com que o processo não permaneça otimizado durante toda a operação.

O uso de otimizadores em tempo real corrobora para correção e permanente condução do processo ao seu ponto ótimo. Neste capítulo serão discutidas as fundamentações teóricas sobre as quais a otimização dinâmica em tempo real está baseada atualmente.

#### 2.1 Princípios da Otimização em Tempo Real

Ao contrário dos algoritmos de controle baseados em otimização, a otimização em tempo real não pode ser baseada em um único tipo de algoritmo. Os controladores baseados em otimização, tais como o MPC (*Model Predictive Controller*) e o NMPC (*Nonlinear Model Predictive Controller*) aplicam-se, de forma geral, a processos contínuos em que o objetivo é manter variáveis controladas (saídas) em um ponto de operação especificado (setpoint) ou em regiões operacionais (faixas) utilizando variáveis capazes de provocar modificações no processo (manipuladas) para compensar os efeitos negativos de variáveis aleatórias ou independentes do processo (distúrbios). Nestes casos, observa-se que a tomada

de decisão depende sempre da distância da predição em relação ao objetivo final, seja ele um *setpoint* ou uma faixa de operação, o que leva o algoritmo de controle a ter uma forma bastante peculiar, e torna o problema de controle passível de ser representado por uma função objetivo fechada e dependente de um modelo de processo, geralmente, bem descrito por um sistema de equações algébrico diferencial (Duraiski 2001) (Duraiski, Trierweiler et al. 2001). Soma-se a isso o fato de que o controle de processo é, de forma geral, feito em pontos de operação estacionários, seja em torno do mesmo ponto ou durante a transição entre dois pontos, o que torna a função objetivo contínua desde que o modelo de equações algébrico-diferenciais que descreve o sistema seja contínuo. Este conjunto de fatores torna a otimização das ações de controle um problema fechado para uma vasta gama de processos variando-se apenas o modelo.

A otimização de um processo, por sua vez, não pode ser representada de forma tão fechada como nos controladores. O primeiro fator que impede a consideração de uma forma fechada diz respeito ao objetivo final da otimização. O que deve ser otimizado? Pelo conceito administrativo de uma empresa um dos principais objetivos é o lucro, mas a questão que é levantada é: O que é o lucro? Se o lucro for determinado como sendo apenas a diferença entre o valor dos produtos e o custo dos insumos, um grave erro pode estar sendo cometido, pois esta é uma visão bastante simplista do ponto de vista financeiro do processo. Além disso, nem sempre máximo lucro de uma unidade produtiva representa o máximo lucro da empresa como um todo. Um exemplo claro desta constatação ocorre na conversão catalítica de gásóleo em produtos combustíveis (Secchi, Santos et al. 2001); (Zanin 2001) onde a função objetivo não é composta de um único termo para o lucro mas sim de uma série de termos que variam de acordo com uma série de fatores. Além da expressão do lucro outras quatro expressões são consideradas como funções objetivo do processo: a máxima conversão de nafta leve (gasolina), a máxima conversão de gás liquefeito de petróleo (GLP), a máxima conversão de óleo leve craqueado (LCO) e a máxima conversão de gásóleo em produtos. Neste processo, estas cinco funções coexistem e são alternadas periodicamente ao longo da operação, de acordo com a programação de produção que trata o problema de otimizar o lucro de uma forma mais geral e do ponto de vista de toda a empresa em conjunto. Isto ocorre devido ao fato de que nem sempre o **lucro** especificado pela função objetivo para a unidade de FCC (Fuel Catalytic Cracking) corresponde ao lucro global da refinaria e da própria empresa como um conjunto de refinarias. Por vezes, a produção de um combustível menos rentável em um dado momento em uma das unidades pode representar um lucro global ótimo para toda a corporação e,

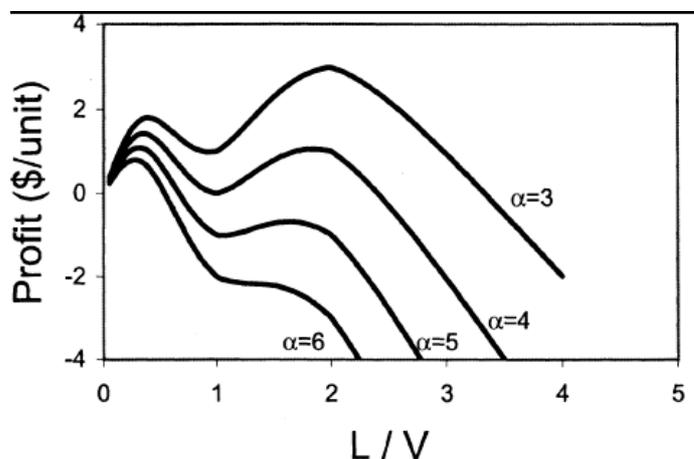
desta forma, o objetivo local da otimização passa a ser a maximização da produção de algum produto específico.

Se por um lado é importante que a otimização em tempo real seja um instrumento bem integrado às camadas superiores do controle dentro do âmbito industrial e deva esta ser uma ferramenta auxiliar da programação de produção, por outro lado, também é importante que a camada de otimização se integre de forma adequada às camadas inferiores do controle (supervisório e regulatório). Não basta que se tome uma estrutura de controle existente em uma unidade e se decida otimizá-la. Sempre é necessário que se questione a validade e a importância da estrutura atual de controle para o alcance dos objetivos aos quais a otimização se propõe. Até que ponto o uso de um *setpoint* em uma variável controlada auxilia na determinação de um ótimo de produção? Até que ponto a utilização de um conjunto específico de variáveis manipuladas é o mais adequado para se manter um processo operando de forma mais lucrativa? A redução da variabilidade torna qualquer processo mais lucrativo ou é melhor permitir que este oscile livremente dentro de uma faixa de especificação? Diversas são as questões que surgem quando se amplia o controle de processos partindo-se do controle supervisório previamente estabelecido e expandindo-o para a camada de otimização, o que faz com que se questione a validade da estrutura atual de controle implementada. Esta constatação é claramente ilustrada em (Loeblein and Perkins 1999) onde um reator de conversão catalítica de combustível é otimizado com duas estruturas diferentes de controle supervisório, o que leva a resultados bastantes distintos em termos de desempenho do otimizador.

Além dos diferentes tipos de denominações do que é o lucro e da integração do otimizador com as camadas superiores e inferiores do controle, uma terceira consideração a ser feita diz respeito ao tipo de otimização a ser realizada. Muitas vezes a otimização aplica-se a processos nos quais todas as variáveis são contínuas tal como em (Zanin 2001) e (Loeblein and Perkins 1999) onde é realizada a otimização de uma unidade de craqueamento catalítico, mas em muitos casos é possível que se tenha otimizações discretas ou mistas. Um exemplo claro de otimização mista é encontrado em (Giovanoglou, Barlatier et al. 2003) onde o objetivo é otimizar a produção de solventes em bateladas. Neste caso, nem todas as misturas de diferentes compostos em diferentes concentrações geram um solventes dentro de especificações aceitáveis pelo mercado. Apenas alguns seletos grupos de combinações de componentes em concentrações específicas apresentam valor comercial de forma que somente estas combinações podem ser aceitas na produção

e as demais são resíduos sem valor. Desta forma, estes grupos de componentes tem que ser representados por variáveis discretas no problema de otimização. Considerando que outras variáveis, tais como temperaturas, fluxos e concentrações dentro de um mesmo grupo, são variáveis contínuas, a otimização deste processo requer um tipo de programação inteira mista.

Uma outra característica da otimização em tempo real é a complexidade relativa a sua abrangência em termos de domínio da função objetivo. Como a otimização se propõe a encontrar o ponto economicamente mais viável de operação de uma unidade, o algoritmo proposto deve observar toda a região possível de operação de modo a encontrar a melhor combinação de variáveis manipuladas, ou seja, o ótimo global. Embora muitos processos apresentem características lineares e permitam a utilização de otimizadores lineares e sem maiores níveis de complexidade, alguns processos apresentam características bastante peculiares e que, se não forem devidamente observadas, podem tornar inútil a aplicação de otimização. Um caso que representa bem esta situação é mostrado em (Lucks 2003) onde se deseja otimizar um processo de absorção de enxofre de óleo cru utilizando-se uma solução aquosa contendo microorganismos.



**Figura 2.1:** Lucro do processo de absorção como função da variável manipulada (relação entre as vazões de óleo e fase aquosa), para diferentes custos da fase aquosa representado pelo parâmetro  $\alpha$ . (Lucks 2003)

Como mostra a Figura 2.1, dependendo do custo da fase aquosa, a curva do lucro da unidade em relação ao custo de tal fase pode assumir duas formas distintas. Na primeira situação, em que o custo da fase aquosa é alto, a curva do lucro mostra um ponto de máximo e um ponto de inflexão. No segundo caso, quando o custo da fase aquosa é reduzido, o ponto de inflexão se transforma em um ponto de máximo, com um valor de lucro maior que o lucro obtido no ponto de máximo fixo. Isto significa que durante uma otimização, se o valor da fase aquosa for elevado o

máximo lucro é alcançado sem maiores problemas por um otimizador comum. No entanto, se o valor da fase aquosa diminuir durante a operação, e o processo estiver otimizado com o valor de fase aquosa alto, uma otimização global será requerida para que o novo ótimo global seja encontrado, uma vez que o ótimo local anterior permanece. Tendo em vista que otimizadores globais requerem um esforço computacional elevado, otimizações deste tipo tornam-se custosas e morosas, restringindo de forma razoável a sua utilização em tempo real. Isto faz com que os otimizadores tenham que funcionar sempre avaliando um compromisso entre o seu desempenho computacional e a qualidade de sua solução.

Uma estratégia complementar, que embora não resolva por completo esse problema, mas que auxilia consideravelmente é a apresentada em (Sequeira 2003), o qual recebe a denominação de Evolução em Tempo Real (RTE – Real Time Evolution). Nesta abordagem, a busca pelo ótimo é abandonada e o objetivo em questão é busca por uma melhora contínua e gradual. Assim sendo, mesmo que o ótimo não seja encontrado no tempo hábil, uma melhora incremental no valor da função objetivo já é o suficiente para que as ações corretivas sejam tomadas no processo. De fato, as constantes melhoras levam o processo ao ótimo de forma gradual ao longo de várias atuações do otimizador. Desta forma, o tempo computacional necessário para que se tenha uma ação aplicável ao processo fica limitada praticamente pelo usuário e o aumento no tempo permitido para a otimização corrobora para um aumento na qualidade da solução o que torna mais rápida a convergência para o ponto ótimo.

## 2.2 Histórico

O conceito de otimizar economicamente o processo iniciou-se de uma forma bastante rudimentar, mas muito efetiva para alguns processos industriais. Neste conceito, a otimização era realizada através de um problema linear (função objetivo e restrições lineares), onde o modelo utilizado representava o ganho de processo em cada variável de. Basicamente, o modelo era constituído de uma matriz de ganhos do sistema (modelo estacionário) e o resultado da otimização eram *setpoints* para as variáveis controladas e *targets* para as variáveis manipuladas, os quais deveriam ser alcançados ao final do período de transição realizado pelo controlador, constituindo um objetivo para o ponto estacionário. Esta técnica levava as variáveis sempre para as suas restrições de mínimo ou de máximo e não era capaz de otimizar processos com variações de ganho. Embora rudimentar, para os recursos computacionais de décadas passadas, esta técnica era bastante eficiente e se aplicava a muitos

processos industriais, sendo ainda o padrão adotado por diversas indústrias dada sua simplicidade e eficácia em situações onde o ótimo operacional encontra-se próximo às restrições operacionais, ou seja, nos casos em que o aumento de produção é sinônimo de aumento de rentabilidade.

No entanto, como dito anteriormente, a otimização linear não tem a capacidade de trabalhar com processos com inversão de ganho, onde os pontos de operação economicamente ótimos não ocorrem sobre as restrições do processo. Por este motivo tornou-se necessário desenvolver uma camada de otimização que não trabalhasse somente com modelos lineares, mas sim com modelos mais completos e não-lineares do processo capazes de descrevê-los com maior precisão o comportamento da unidade industrial. Surgiu assim o conceito de RTO (Real-Time Optimization). Junto com o seu surgimento, um problema inerente à sua aplicação prática teve que ser resolvido para que este fosse efetivamente aplicado - a ausência de retro-alimentação do processo. A camada de otimização, quando linear, não precisa da retro-alimentação do processo uma vez que trabalha com os valores relativos, ou desvios, e o ótimo sempre ocorre sobre as restrições. Porém, quando a otimização passa a ser não-linear, as variáveis são consideradas em seus valores absolutos, ou posicionais, e é necessário que o otimizador “se localize” dentro do processo para que possa realmente movê-lo para o ponto ótimo. A forma encontrada para se fazer esta retro-alimentação dos dados do processo foi a realização de uma reconciliação de dados (Zhang and Forbes 2000). Uma forma de considerar a estimação dos parâmetros, juntamente com os distúrbios medidos do processo é mostrada em (Dordelly and Marhn 2000). Nesta abordagem, a função objetivo é modificada acrescentando termos para considerar estes distúrbios. Trata-se de uma forma de função objetivo para englobar incertezas do modelo (parâmetros) e distúrbios medidos. O primeiro termo da função objetivo é a expressão do lucro do processo o qual é seguido por uma série de aproximações quadráticas (expansão em série de Taylor truncada no termo quadrático) da expressão do lucro e tendo como variáveis os parâmetros e os distúrbios medidos. Esta aproximação descreve a sensibilidade da função objetivo em relação a cada variável.

O inconveniente nas otimizações em tempo real, sejam elas lineares ou não-lineares, é que sua execução necessita que o processo esteja em um estado estacionário, ou, ao menos, tenha uma predição deste, para que os resultados sejam consistentes. A detecção de estados estacionários de processo não é trivial e várias técnicas de detecção de estacionários têm sido desenvolvidas como mostra (Jiang, Chen et al. 2003) e (Caumo 2006). Nestes trabalhos, a detecção do estado

estacionário é realizada utilizando-se *wavelets*, uma espécie de extensão da transformada de Fourier que considera janelas temporais a fim de verificar a distribuição de frequências do sinal considerando o instante de tempo em que as frequências se manifestam. Através desta decomposição são gerados filtros do sinal do processo e através de critérios de tolerância é possível determinar quando o processo encontra-se no estado estacionário.

A evolução em termos de otimização de processos fez com que a determinação dos ótimos econômicos considerasse problemas dinâmicos, onde os *targets* e *setpoints* do processo passaram a ser determinados para intervalos de tempo de predição do ótimo. Esta evolução permitiu fazer a otimização não só buscar o melhor ponto de operação econômico como também determinar a forma mais econômica de se fazer esta transição. Este novo advento torna-se interessante nos casos em que se tem, por exemplo, duas variáveis manipuladas capazes de fazer uma mudança de ponto de operação do processo, uma de resposta muito rápida e de custo mais elevado e outra de custo mais baixo e mais lenta. Um otimizador estacionário modificaria a variável de custo mais baixo para se chegar ao ponto de operação desejado uma vez que estaria vendo apenas o ganho estacionário do processo. Um otimizador dinâmico movimentaria inicialmente a variável mais rápida e mais cara para fazer com que o processo chegasse ao ponto de operação o mais breve possível e depois retornaria com esta variável lentamente, compensando com a variável de custo mais baixo.

Além desta característica, a utilização de modelos dinâmicos permitiu o desenvolvimento de camadas de otimização acoplada a camada de controle supervísório (Tvrzka and Oadloak 1998). A forma até então utilizada para a aplicação de otimização em tempo real consistia em executar um algoritmo de otimização que fornecia *targets* e *setpoints* para um algoritmo de controle que, por sua vez, determinava as ações de controle para o processo. Esta abordagem é conhecida como estrutura multicamada na qual o otimizador encontra-se em uma camada superior ao controle supervísório. A abordagem proposta em (Tvrzka and Oadloak 1998) é a denominada monocamada onde o otimizador e controlador supervísório trabalham em um único algoritmo. Embora esta idéia seja muito atrativa, ela tem como principal desvantagem a sua maior suscetibilidade a falhas, a qual é crítica quando se trabalha com processos industriais.

## 2.3 Algoritmos de Otimização

A estratégia utilizada nesta tese para a solução do problema algébrico de otimização consiste na transformação do problema original de otimização em um problema com função objetivo linear e restrições não-lineares. Assim, afora os casos em que as características do problema permitem uma solução totalmente particularizada como em (Cheraghi, Lim et al. 1996), a literatura sugere a utilização de um método SLP (Sequential Linear Programming) ou ILP (Iterated Linear Programming) como apresentado em (Shyamsundar and Rangaiah 2000), (Han and Rangaiah 1997), (Lim and Cheraghi 1998), (Gopal and Biegler 1997) e (Syam 2005).

É consideravelmente grande a quantidade de algoritmos desenvolvidos para resolver problemas de otimização numérica e seria pouco útil fazer menção a todos neste trabalho. Os métodos mais significativos e relacionados com o desenvolvimento deste trabalho são a programação linear, programação linear seqüencial, programação quadrática, programação quadrática seqüencial e método dos gradientes reduzidos. Devido à consagração destes métodos pela literatura, os mesmos não serão detalhados nesta revisão bibliográfica. Maiores informações podem ser encontradas em (Nocedal and Wright 1999).

## 2.4 Algoritmos Aplicados a RTO e DRTO (Dynamic Real-Time Optimization)

Devido às características previamente expostas nas seções anteriores, os problemas de otimização em tempo real de processos, sejam eles dinâmicos ou estacionários possuem características bastante particulares e que exigem, na sua aplicação, o estabelecimento de adaptações e combinações de algoritmos para sua realização.

A otimização, quando realizada de forma dinâmica, necessita de dois tipos de algoritmos associados: um para a resolução do problema de otimização propriamente dito e outro para a resolução do sistema de equações algébrico-diferenciais que surge da modelagem dinâmica do processo industrial. De forma geral, estes algoritmos podem ser classificados em dois grandes grupos. Um deles, denominado de algoritmos seqüenciais, é o que resolve os dois problemas de forma alternada e iterativa até a obtenção da solução. O outro grupo, denominado de algoritmos simultâneos, é o dos algoritmos que acoplam os dois problemas em uma única etapa de resolução.

A abordagem de (Biegler, Cervantes et al. 2002) apresenta um algoritmo simultâneo de resolução do problema de otimização dinâmica em tempo real. Neste algoritmo, a programação não-linear é baseada no método do ponto interior. A direção de busca é determinada pelo método dos gradientes conjugados. O sistema de equações algébrico-diferenciais é discretizado utilizando o método de colocação em elementos finitos.

A abordagem adotada por (Diehl 2001) e (Diehl, Bock et al. 2002) é uma mistura entre as abordagens seqüencial e simultânea, denominada de multi-shooting. Neste caso o intervalo de integração é particionado e múltiplos pontos iniciais são distribuídos sobre uma primeira trajetória hipotética. Assim, tais pontos denominados nodos são utilizados como graus de liberdade do problema de otimização e restrições de igualdade obrigam os pontos finais de cada uma das trajetórias geradas a partir de cada um desses nodos a coincidirem com o seguinte. O problema originalmente proposto:

$$J = \min_u \sum_{i=0}^N g(x_i, u)$$

onde :

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

$$x(t_0) = x_0$$
(2.1)

É transformado no problema:

$$J = \min_{u,s} \sum_{i=0}^N g(s_i, u)$$

sujeito as restrições :

$$\frac{dx}{dt} = f(t, x, u), \quad t \in [t_i, t_{i+1}]$$

$$s_{i+1} - x_i(t_{i+1}) = 0 \quad \text{para } i = 0, 1, \dots, N-1$$

$$s_0 - x_0(t_0) = 0$$
(2.2)

Este método é dito ser intermediário às abordagens seqüencial e simultânea por ser capaz de se aproximar de ambas de acordo com o intervalo de discretização utilizado. Quando os intervalos são suficientemente pequenos e a quantidade de nodos distribuídos for relativamente grande, tem-se um método totalmente simultâneo, onde a integração é eliminada e todo o problema é resolvido em uma única etapa. Por outro lado, quando o intervalo de discretização for do tamanho do

horizonte de integração do modelo de equações algébrico-diferenciais, tem-se o método totalmente seqüencial onde em cada iteração do otimizador tem-se a resolução do sistema de equações algébrico-diferenciais.

## 2.5 Modelos para Otimização

A otimização em tempo real, nos padrões atuais, pressupõe a representação do processo na forma de um modelo não-linear, capaz de descrever todas as relações entre as variáveis, inversões de ganho e mudança de dominâncias de variáveis no processo. Um modelo com estas características é, geralmente, descrito por um sistema de equações algébrico-diferenciais, que poderia incluir equações diferenciais parciais.

Devido à complexidade do cálculo e da dificuldade de se encontrar um ótimo utilizando este modelo, procura-se, de forma geral, simplificar o modelo a fim de estabelecer um compromisso entre a reprodutibilidade do processo real e o custo computacional para o seu cálculo.

A primeira simplificação bastante comum e que não leva a uma perda de qualidade significativa é a discretização das variáveis espaciais, transformando equações diferenciais parciais em equações diferenciais ordinárias, também conhecido como aplicação do método das linhas (Nocedal and Wright 1999).

Ainda assim, modelos fenomenológicos descritos através de sistemas de equações algébrico-diferenciais são bastante custosos em dois sentidos: um por apresentarem um alto custo computacional para a sua simulação, a exemplo de colunas de destilação que podem ter centenas de estados, e outro por levarem um tempo considerável para serem construídos e validados. Um modelo fenomenológico completo pode levar meses para ser concluído, o que no âmbito industrial nem sempre é permitido.

Para reduzir o esforço computacional na simulação de modelos tão complexos, técnicas de redução de ordem de sistemas de equações algébrico-diferenciais, tais como as apresentados em (Hedengren and Edgar 2005), podem ser aplicadas, no entanto, a dificuldade de obtenção destes modelos continua sendo a mesma.

Alternativas aos modelos fenomenológicos podem ser encontradas na literatura, juntamente com as suas metodologias de identificação não-linear e que são estudadas neste trabalho.

A forma abordada em (Fernandes 2006) mostra uma metodologia de identificação de modelos não-lineares baseada na interpolação de modelos lineares identificado em diferentes regiões de operação. A técnica consiste em determinar um sistema autônomo dado pela expressão:

$$\frac{dx}{dt} = A(u) \cdot (x - \Omega(u)) \quad (2.3)$$

Neste sistema autônomo,  $A(u)$  é a interpolação das matrizes  $A$  do espaço de estado de cada um dos modelos lineares identificado em cada uma das regiões de operação, usando como variável independente para a interpolação os valores das variáveis de entrada  $u$ . O vetor  $\Omega(u)$  corresponde aos valores sem serem expressos como variáveis desvio, ou seja, na sua forma original, dos estados estacionários  $x$  de cada um dos modelos lineares identificados, interpolados também em função das variáveis  $u$  de entrada. O vetor  $x$  corresponde aos estados atuais do modelo.

Esta abordagem apresenta, no entanto, alguns problemas práticos na sua implementação. O primeiro problema diz respeito à obtenção dos modelos lineares com ordens iguais. Como a interpolação da matriz  $A$  é feita termo a termo, é necessário que todas as matrizes  $A$  de todos os modelos lineares identificados tenham as mesmas dimensões. Isto significa que todos os modelos devem conter o mesmo número de estados, ou em outras palavras, possuir a mesma ordem, o que nem sempre é possível uma vez que para diferentes regiões de operação é possível representar o processo através de modelos lineares com diferentes ordens dependendo da complexidade da sua resposta. Além disso, mesmo que se obtenham modelos com ordens iguais nada garante que cada um dos estados conserve uma equivalência em todos os modelos de forma que seja possível interpolá-los de forma lógica. Para exemplificar pode-se citar o caso de um modelo que tenha os estados  $x_1$  e  $x_2$  no ponto de operação A e os estados  $z_1$  e  $z_2$  no ponto de operação B. Nada garante que se deva interpolar os estados  $x_1$  e  $z_1$  em uma função e os estados  $x_2$  e  $z_2$  em outra, pois nada indica que exista uma equivalência direta entre estes estados. Poderia até mesmo ser mais conveniente interpolar  $x_1$  e  $z_2$  em uma função e  $x_2$  e  $z_1$  em lugar de  $x_1$ - $z_1$  e  $x_2$ - $z_2$ . Embora para dois estados seja fácil determinar qual a melhor relação, para casos onde o número de estados começa a se tornar elevado esta análise tende a ser inviável. Para amenizar este

problema pode-se recorrer à alternativa de reduzir-se cada sistema a uma base canônica, mas esta alternativa nem sempre dá bons resultados.

Um terceiro problema que ocorre neste tipo de interpolação de modelos está na forma como as matrizes  $A$  são interpoladas. Ainda que se tenha o mesmo número de estados e que estes conservem uma relação direta entre eles, a interpolação de valores da matriz  $A$  pode gerar funções que, para certos valores de variáveis manipuladas, seus valores característicos sejam positivos. Isto leva a geração de sistemas instáveis e que não descrevem e nem se aproximam do comportamento do processo nestas regiões intermediárias.

Todas estas limitações referentes à interpolação de pólos e estados estacionários do sistema levaram a uma evolução do método mostrada em (Duraiski and Trierweiler 2005). Neste método as respostas do sistema são consideradas por canais onde cada canal possui um conjunto de estados associados de acordo com a ordem da resposta. A cada estado é associado um pólo de forma que as respostas de cada canal são constituídas por uma combinação de respostas de primeira ordem. Desta forma a interpolação acontece diretamente sobre o pólo o que torna mais fácil garantir a sua negatividade. Além disso, nos casos de redução de ordem em um dado canal não são tratados como uma redução de estados, mas sim a redução do ganho a zero para um estado. Um detalhamento mais completo deste método que é a base da metodologia de identificação de modelos proposto nesta tese será discutido em detalhes no Capítulo 4.

## 2.6 A Interação Otimizador-Controlador

Para que a otimização faça surtir no processo os efeitos desejados, é necessário que toda a cadeia de dispositivos que atuam no processo esteja funcionando de forma integrada e ajustada para permitir um funcionamento adequado destes dispositivos. Assim sendo, é necessário que os controladores regulatórios da planta estejam ajustados para que as mudanças de *setpoint* enviadas pelo controle supervisão sejam capazes de serem realizadas de forma eficaz. De forma análoga, quando se ajusta o controle supervisão existe uma diferença considerável no seu ajuste para os casos onde se aplicará uma camada de otimização superior ou não. Quando o controlador é ajustado para trabalhar sem a camada de otimização, os únicos ajustes necessários são os pesos do erro, supressão de movimentos, horizonte de predição, horizonte de controle, e tempo de amostragem ou *blocking* (dependendo da metodologia do controlador). No caso em que uma camada de otimização superior é acoplada, surge um novo parâmetro de

ajuste para o controlador preditivo que é o peso do *target*. Os *targets* são valores de operação especificados para as variáveis manipuladas e que, de forma geral, não são utilizados quando não se trabalha com a otimização do sistema, dispensando assim um ajuste para o seu peso. No entanto, quando a camada de otimização é ativada e têm-se mais variáveis manipuladas do que controladas, os *targets* são especificados para estas variáveis manipuladas excedentes e o ajuste correto de seus pesos é fundamental para que o controlador preditivo não perca o seu desempenho. Para agravar a situação, os pesos dos *targets* são relativos aos demais pesos especificados para os outros termos da função objetivo, o que faz com que o ajuste de todos os pesos deva ser feito em uma única etapa.

Uma metodologia para determinar esses pesos de forma conjunta é apresentada em (Kassidas, Patry et al. 2000), onde uma otimização é feita para que se determine os pesos, *setpoints* e *targets* para o controlador. A estratégia consiste em otimizar, além dos *targets* e dos *setpoints*, os pesos do erro e do *target* em um único problema de otimização. No referido trabalho, a otimização é realizada *off-line* como uma etapa de projeto do controlador preditivo. O modelo é construído de tal forma que considera como variáveis de entrada as manipuláveis e os distúrbios do processo, e, para determinar o ponto ótimo de forma *off-line*, uma seqüência de distúrbios típicos é realizada e os parâmetros (pesos, *setpoints* e *targets*) são ajustados para estes distúrbios típicos. Embora toda a consideração aqui seja feita de forma *off-line*, o estudo do comportamento desta forma de otimizar pode ser interessante.

## 2.7 Tratamento de Incertezas

Embora se tenha uma representação do processo através de modelos e se procure adaptá-lo da melhor forma possível às respostas do sistema, sempre haverá incertezas inerentes a diversos fatores na relação dos sistemas de controle com o processo. Segundo (Zhang, Nadler et al. 2001), as incertezas na aplicação de um sistema de RTO podem ser de 4 tipos: incertezas de mercado, incertezas de processo, incertezas de medida e incertezas de modelagem, as quais são discutidas nas seções subseqüentes

### 2.7.1 Incertezas de mercado

Incertezas de mercado são aquelas que se referem diretamente às condições econômicas ao qual o processo está sujeito. Nem sempre é fácil avaliar o custo de insumos, bem como não é completamente definido o valor que os produtos têm no

mercado, sendo que todos estes valores tratam-se de estimativas e valores médios, de forma geral. Além disso, a modelagem de custo e lucro estabelecido pelas funções objetivos usualmente utilizadas são definições extremamente simplistas destas grandezas. Se a definição de lucro fosse feita de um ponto de vista mais administrativo do processo, uma série de fatores deveria ser incluída nesta determinação. Além dos custos dos insumos e preços de venda dos produtos, dois importantes grupos de custos fazem parte do custo real de um produto e que são completamente desconsiderados na prática: os custos fixos e administrativos.

Os custos fixos são aqueles que, embora não esteja havendo produção sempre incidirão sobre o processo. Enquadram-se nestes custos, pessoal de operação, aluguel de equipamentos, etc. Estes custos impactam diretamente no valor do produto e reduzem o lucro total, mas são constantes, independente do volume da produção, de forma que quanto maior for a produtividade mais diluído será a sua interferência.

Os custos administrativos, por sua vez, são similares aos custos fixos com a principal diferença que estes estão inseridos em um âmbito mais global da unidade produtiva e dizem respeito a vários processos. As indústrias costumam ter vários processos produtivos que levam a diferentes produtos. Ao nível gerencial todos são administrados em conjunto de forma que não existe uma clareza quanto à parcela da folha de pagamento dos funcionários do setor administrativo que é paga por um produto específico, e mesmo que estimativas possam ser feitas a respeito desses custos, sempre serão estimativas por mais precisas que sejam. Exemplos claros destes custos são impostos territoriais, pessoal administrativo, despesas de telefone e material de consumo em escritórios, etc..

Ainda que se englobassem na função objetivo tais custos, as formas de predizê-los poderia ser tão complexa que jamais poderiam ser avaliados com precisão.

### **2.7.2 Incertezas de processo**

Diferentemente das incertezas de mercado, as incertezas de processo fogem do universo global no qual o processo está inserido e abrangem um grupo mais restrito de variáveis. Estas incertezas derivam-se da incapacidade de se modelar com exatidão muitos fenômenos físicos que influenciam no processo. Um exemplo bastante claro deste tipo de incerteza é a temperatura externa (ambiente) de uma coluna de destilação. Por melhor que seja o isolamento térmico deste equipamento,

sempre haverá uma troca térmica entre o fluido que circula na coluna e meio ambiente, a qual é extremamente dependente da temperatura do ambiente no qual a mesma está inserida. A quantidade de fatores que influenciam nesta temperatura é tão grande que nem mesmo o melhor modelo climático poderia descrevê-la com precisão. Ainda que esta temperatura fosse medida, a troca térmica pode ser influenciada por tantos fatores climáticos, tais como chuva, ventos, umidade do ar, irradiação solar entre outros de forma que seria difícil precisar a troca térmica realmente realizada.

Além dos fatores ambientais, existe ainda uma série de fatores que não podem ser modelados no processo devido a sua complexidade ou falta de conhecimento do fenômeno. Um exemplo claro disto são as atribuições de fatores de eficiência às colunas de destilação. Esta atribuição vem do fato que se assumem, geralmente, modelos em equilíbrio termodinâmico de tal forma que os efeitos de transporte de energia, massa e quantidade de movimento não são modelados. As trocas de massa nos pratos recebem uma descrição deficiente do fenômeno, o que obriga a se fazer a consideração de um fator de eficiência que procura corrigir todos estes fenômenos físicos não modelados. Por outro lado, se a modelagem fosse feita de uma forma mais exata, a complexidade de sua resolução seria tal que tornaria impraticável a utilização deste modelo para os recursos computacionais e métodos científicos de resolução dos tempos atuais.

Este conjunto de fatores leva a uma “descrição aproximada” da realidade o que fica subentendido na palavra modelo do processo.

### **2.7.3 Incertezas de medida**

Somado aos fatores anteriores, as incertezas de medida contribuem para a falta de uma garantia do que está realmente acontecendo com o processo industrial. As incertezas de medida são provocadas pela falta de capacidade dos instrumentos que fornecem as informações do processo traduzirem a realidade.

Um dos fatores que fazem com que esta medida seja imprecisa é o ruído. O ruído ocorre devido ao fato de que todo o instrumento de medida esta sujeito a fenômenos, além do qual está designado a medir, que não podem ser isolados ou removidos de sua metodologia de aquisição e determinação da medida, e por isso sempre estarão presentes em menor ou maior amplitude. Somado a existência de ruído, o surgimento de *bias* referente à perda de calibração durante o uso do equipamento também é algo muito comum e dificilmente pode ser evitado.

### **2.7.4 Incertezas de modelo**

Por fim, por melhor que seja o conhecimento sobre as leis físicas que descrevem um processo industrial, o desenvolvimento de um modelo está intimamente relacionado ao surgimento de parâmetros, os quais precisam ter seus valores ajustados para descreverem o processo. Este ajuste depende de testes experimentais e perturbações no processo e estão muito susceptíveis a falhas ou imprecisões, por vários motivos. Um deles é o fato de que testes de bancada não são testes em processo, e embora estes conservem uma boa reprodutibilidade, as condições em que os fenômenos ocorrem em bancada e em processo são diferentes e isso gera imprecisão.

Mesmo que tais testes sejam realizados em planta, fenômenos escusos à vontade do pesquisador podem tornar os dados ligeiramente divergentes da realidade. Soma-se a isso a inconstância dos processos industriais no tempo, tal como a degradação de catalisadores, incrustação em tubos de trocadores de calor e outros fenômenos que fazem com que parâmetros levantados percam a sua reprodutibilidade com o tempo.

### **2.7.5 Métodos para trabalhar com incertezas**

Quando um sistema de otimização em tempo real é aplicado, além da otimização propriamente dita, alguns passos anteriores são realizados para se prevenir que algumas destas incertezas tornem a otimização inútil ou não aplicável. A seqüência de passos para a localização do ponto ótimo inicia-se com uma reconciliação de dados. Nesta reconciliação, procura-se de forma ótima satisfazer os balanços de massa e energia do processo a fim de se certificar do valor correto das medidas, o que reduz consideravelmente a influência das incertezas de medida do processo. Logo após, de posse das medidas “supostamente” corretas do processo, estes dados são utilizados para se fazer uma estimativa de parâmetros do modelo do processo, o que minimiza a ocorrência de incertezas de modelo do processo. Algumas formas de se fazer estas estimativas de parâmetros são descritas em (Duraiski and Trierweiler 2005) e (Shouche, Genceli et al. 2002). Posteriormente, a otimização do processo é realizada, sem que as incertezas de processo e econômicas tenham sido consideradas.

Estas incertezas podem levar a otimização em tempo real a encontrar ótimos em regiões não viáveis na prática, o que faz com que uma otimização tenha que ser desconsiderada em um dado momento. Por este motivo, a literatura cita métodos

para trabalhar com estas incertezas de forma a quantificá-las com uma melhor aproximação e, conseqüentemente, tornando a otimização mais robusta.

Em (Zhang, Monder et al. 2002), um método estatístico de consideração das incertezas é apresentado. Neste método, duas modificações importantes são realizadas no problema de otimização. Dado o problema de otimização:

$$\begin{aligned} \min_x & (J(x, \theta, \varphi)) \\ \text{sujeito a} & : g(x, \varphi) \geq 0 \end{aligned} \quad (2.4)$$

onde  $x$  são as variáveis independentes do processo;  $\theta$  corresponde à incerteza nos parâmetros de mercado (existente apenas na função objetivo);  $\varphi$  é a incerteza nos parâmetros de processo;  $J$  é a função objetivo que engloba as restrições de igualdade do modelo; e  $g$  as restrições de desigualdade do problema. Para a consideração estatística do problema, as incertezas  $\theta$  e  $\varphi$  são consideradas como distribuições normais e a função objetivo é considerada como a esperança da função objetivo original somada ao seu desvio padrão, da forma:

$$\begin{aligned} \min_x & (E(J(x, \theta(\varphi))) + \sigma(J(x, \theta, \varphi))) \\ \text{sujeito a} & : g(x, \varphi) \geq 0 \end{aligned}$$

A formulação acima é uma descrição rigorosa da consideração estatística da função objetivo. No entanto, algumas simplificações podem ser feitas devido às características do problema. Considerando que os desvios padrão das incertezas não variam muito para os diferentes pontos de operação do sistema este termo pode ser considerado constante e desprezado da função objetivo.

A restrição de desigualdade, por sua vez, passa a ser considerada como a esperança do valor da função  $g$  maior que zero ou a esperança da função  $g$  negativa para  $g$  menor que zero seja menor que um desvio  $d$  estabelecido, ou seja:

$$\begin{aligned} \min_x & (E(J(x, \theta(\varphi))) + \sigma(J(x, \theta, \varphi))) \\ \text{sujeito a} & : E(-g(x, \varphi) | g(x, \varphi) < 0) \leq d \end{aligned} \quad (2.5)$$

A consideração probabilística da função objetivo, embora não seja indispensável, torna a resolução do problema de otimização mais robusta e menos susceptível a falhas segundo (Zhang, Monder et al. 2002).

## 2.8 Referências Bibliográficas

- Biegler, L. T., A. M. Cervantes, et al. (2002). "Advances in Simultaneous Strategies for Dynamic Process Optimization." Chemical Engineering Science **57**: 575 - 595.
- Caumo, L. (2006). Departamento de Engenharia Química - Escola de Engenharia. Porto Alegre, Universidade Federal do Rio Grande do Sul. Mestrado.
- Cheraghi, S. H., H. S. Lim, et al. (1996). "Straightness and Flatness Tolerance Evaluation: An Optimization Approach." Precision Engineering **18**: 30 - 37.
- Diehl, M. M. (2001). Real Time Optimization for Large Scale Nonlinear Process. Heidelberg, Ruprecht-karis-Universität. **Doutorado**: 201.
- Diehl, M. M., H. G. Bock, et al. (2002). "Real-Time Optimization And Nonlinear Model Predictive Control Of Processes Governed By Differential-Algebraic Equations." Journal of Process Control **12**: 577 - 585.
- Dordelly, J. L. C. and T. E. Marhn (2000). "Control Design for Increased Profit." Computers and Chemical Engineering **24**: 267 - 272.
- Duraiski, R. G. (2001). Controle Preditivo Não-Linear Utilizando Linearizações ao Longo da Trajetória. Departamento de Engenharia Química - Escola de Engenharia. Porto Alegre, Universidade Federal do Rio Grande do Sul. Mestrado: 112.
- Duraiski, R. G. and J. O. Trierweiler (2005). A Novel Gray-Box Nonlinear Dynamic Model Identification Algorithm Based On The Multishooting Approach. Encontro de Engenharia de Processo do Mercosur (ENPROMER), Mangaratiba.
- Duraiski, R. G., J. O. Trierweiler, et al. (2001). Nonlinear Model Predictive Control Using Successive Linearization Approach. ENPROMER - Encontro de Engenharia de Processo do Mercosul, Santa Fé - Argentina.
- Fernandes, P. R. B. (2006). The Input-Parameterized Linearization around the Equilibrium Manifold Approach to Modeling and Identification. Dortmund, Universidade de Dortmund. **Doutorado**: 180.
- Giovanoglou, A., J. Barlatier, et al. (2003). "Optimal Solvent Design for Batch Separation Based on Economic Performance." AIChE Journal **49**(12): 3095 - 3109.
- Gopal, V. and L. T. Biegler (1997). "Nonsmooth Dynamic Simulation with Linear Program Based Method." Computers and Chemical Engineering **21**(7): 675 - 689.
- Han, G. and G. P. Rangaiah (1997). "A method for Calculation vapour-liquid and liquid-liquid equilibrium." Computers and Chemical Engineering **21**: 905 - 913.
- Hedengren, J. D. and T. F. Edgar (2005). Order Reduction Of Large Scale DAE Models. Symposium on Modeling of Complex Process, Texas.
- Jiang, T., B. Chen, et al. (2003). "Application Of Steady-State Detection Method Based On Wavelet Transform." Computers and Chemical Engineering **27**: 569 - 578.

- Kassidas, A., J. Patry, et al. (2000). "Integrating Process and Controller Models for the Design of Self-Optimizing Control." Computers and Chemical Engineering **24**: 2589 - 2602.
- Lim, H. S. and S. H. Cheraghi (1998). "An Optimization Approach to Shape Matching and Recognition." Computers and Chemical Engineering **24**: 183 - 200.
- Loeblein, C. and J. D. Perkins (1999). "Structural Design for On-Line Process Optimization: II. Application to a Simulated FCC." AIChE Journal **45**(5): 1030 - 1040.
- Lucks, D. (2003). "Real-Time Optimization in Nonlinear Chemical Processes: Need For Global Optimizer." AIChE Journal **49**: 2980 - 2983.
- Nocedal, J. and S. Wright (1999). Numerical Optimization, Springer Verlag.
- Secchi, A. R., M. G. Santos, et al. (2001). "A Dynamic Model for a FCC UOP Stacked Converter Unit." Computers and Chemical Engineering **25**: 851 - 858.
- Sequeira, S. E. (2003). Real Time Evolution (RTE) for on-line optimisation of continuous and semi-continuous chemical processes. Departament d'Enginyeria Química - Escola Tècnica Superior d'Enginyeria Industrial de Barcelona. Barcelona, Universitat Politècnica de Catalunya. **Doutorado**: 201.
- Shouche, M. S., H. Genceli, et al. (2002). "Effect Of On-Line Optimization Techniques On Model Predictive Control And Identification (Mpci)." Computers and Chemical Engineering **26**: 1241 - 1252.
- Shyamsundar, V. and G. P. Rangaiah (2000). "A Method for Simulation and Optimization of Multiphase Distillation." Computers and Chemical Engineering **24**: 23 - 37.
- Syam, M. I. (2005). "Nonlinear Optimization Exclusion Tests for Finding All Solutions of Nonlinear Equations." Applied Mathematics and Computation **170**: 4104 - 4116.
- Tvrzcka, M. and D. Oadloak (1998). "One layer Real Time Optimization of LPG Production in the FCC Unit Procedure, Advantages and Disadvantages." Computers and Chemical Engineering **22**: 191 - 205.
- Zanin, A. c. (2001). implementação industrial de um Otimizador em Tempo Real. Escola Politécnica - Departamento de Engenharia Química. São Paulo, Universidade de São Paulo. **Doutorado**: 183.
- Zhang, Y. and J. F. Forbes (2000). "Extended Design Cost: A Performance Criterion For Real-Time Optimization Systems." Computers and Chemical Engineering **24**: 1829 - 1841.
- Zhang, Y., D. Monder, et al. (2002). "Real-Time Optimization Under Parametric Uncertainty::A Probability Constrained Approach." Journal of Process Control **12**: 373 - 389.
- Zhang, Y., D. Nadler, et al. (2001). "Results Analysis for Trust Constrained Real-Time Optimization." Journal of Process Control **11**: 329 - 341.

## Capítulo 3

### Algoritmo de Otimização GA-GRG

Neste capítulo, será apresentada a proposta de um novo algoritmo baseado no método de perseguição dos objetivos (*goal attainment*). O algoritmo, denominado GA-GRG (*Goal Attainment based on the Generalized Reduced Gradient*), é um método mais apropriado para resolver problemas de otimização com múltiplos objetivos. Este algoritmo foi desenvolvido baseado no método dos Gradientes Reduzidos Generalizados, com importantes modificações, tendo como objetivo melhorar a convergência para a solução.

A primeira modificação importante consiste em executar uma fatoração do gradiente das restrições, o que permite encontrar as variáveis independentes mais sensíveis para a direção de busca, evitando escolher variáveis inadequadas. Outra modificação importante é o modo pelo qual a direção de busca é determinada. Neste algoritmo, determina-se a direção de busca através de um problema de otimização com solução analítica aplicável no algoritmo proposto, para qualquer problema, independente da não linearidade das restrições e das funções objetivo. A terceira modificação importante deste algoritmo consiste em mudar o problema de otimização não-linear original composto por um conjunto de restrições e funções objetivo não-lineares em um problema com uma função objetivo linear sujeito a restrições não-lineares, que obrigam a solução a caminhar sobre as restrições (método dos conjuntos ativos fixo) do sistema modificado. Os resultados mostram um desempenho alto do algoritmo quando aplicado a vários problemas de

otimização. O Apêndice A traz na íntegra o artigo submetido à publicação, o qual descreve o método e seus resultados.

## **Apêndice A**

**Artigo: “GA-GRG Optimization: A New Algorithm For Multi-Objective Problems Based on the Goal Attainment Method”**



## **GA-GRG Optimization: A New Algorithm for Multi-Objective Problems Based on the Goal Attainment Method**

**Ricardo G. Duraïski · Jorge O. Trierweiler · Argimiro R. Secchi**

**Abstract:** A new algorithm based on the goal attainment method is proposed. The algorithm, named GA-GRG, is a more appropriate way to settle multi-objective optimization problems. The algorithm is also based on the GRG method, with important changes to improve the convergence to the solution. The first modification consists of performing a factorization of the gradient of the constraints, allowing finding the most sensitive independent variables to the search direction, avoiding choosing inappropriate variables. Other important modification is the way in which the search direction is determined. In the proposed algorithm, the search direction is determined through an optimization problem with analytic solution applicable to any problem, independent on the nonlinearities of the constraint and objective functions. The third important modification of this algorithm consists of changing the original nonlinear problem composed by a set of nonlinear objective functions and nonlinear constraints into a problem with one linear objective function subject to nonlinear constraints, which allows the algorithm to ride only on the constraints (Active Set Method) of the modified system. The results show the high performance and high accuracy of the algorithm when applied to several optimization problems.

**Keywords:** Multi-objective optimization, Goal Attainment, Optimization Algorithm.

---

Ricardo G. Duraïski · Jorge O. Trierweiler  
Group of Integration, Modeling, Simulation, Control and Optimization of Processes (GIMSCOP)  
Chemical Engineering Department - Federal University of Rio Grande do Sul (UFRGS)  
Rua Eng. Luiz Englert s/n – CEP: 90040-040 – Porto Alegre, RS – Brazil  
Phone: +55-51-3308-4072 – Fax: +55-51-3308-3277  
E-mail: {rduraïsk, jorge}@enq.ufrgs.br

Argimiro R. Secchi (corresponding author)  
COPPE/PEQ/LMSCP – Federal University of Rio de Janeiro (UFRJ)  
Av. Horácio Macedo, 2030 - Centro de Tecnologia. CEP 21941-914, Rio de Janeiro, RJ - Brazil.  
Phone: +55-21-2562-8301 – Fax: +55-21-2562-8300  
E-mail: arge@peq.coppe.ufrj.br

## 1 Introduction

Multi-objective algorithms have been extensively studied for many researchers as a powerful tool for settling engineering problems. However, the largest amount of papers found in literature about this kind of problems is applied to electrical and computer engineering. For instance in (Kleemann et al. 2007) the multi-objective problem was proposed to design a Laser transmitter. In (Wang et al. 2006) a multi-objective problem was settled to improve the QoS (Quality of Service) of a multicasting computer network. In the same way (Aboeela and Douligeris 1999) uses the multi-objective optimization for optimizing the QoS of touting internet systems. In (Roth et al. 2006) and (Chiam et al. 2007), the multi-objective optimization is used to estimate structures of neural networks for pattern recognition and time series generation. The special characteristic of the last one is the usage of a fuzzy objective function which is an innovation in this area.

As it can be seen, the multi-objective optimization is extensively used in problems with searches for the global minimum and with discrete variables. Probably for these reasons, the main kind of algorithm used for these optimizations is the stochastic genetic evolutive algorithms (Mezura-Montes and Coello 2004), (Coello 1998). Improvements of this algorithm have been suggested in the literature. In (Parks and Miller 1998), it is shown that the increasing of elitism and selection pressure can improve considerably the performance of the algorithm. (Kukkonen and Lampinen 2004) suggests the Generalized Differential Evolutive Genetic Algorithm (GDE) which is a genetic algorithm with a special characteristic. There is an upper limit for the number of generations in this algorithm and the modifications performed allow the consideration of constrains. In (Sareni et al. 2004), the improvement of the genetic algorithm is made by modifying the steps of recombination and self adaptation. In the field of algorithms used for multi-objective optimization, it can be also mentioned the algorithms MOMGA-IIa (multi-objective messy genetic algorithm) (Day and Lamont 2005) and the algorithm PCGA II (Pareto Constrained Genetic Algorithm) (Kumar et al. 2004).

Considering that the intention here is to use optimization in real time, a new algorithm, specially designed for goal attainment has been developed. The GA-GRG (Goal Attainment Generalized Reduced Gradients) is an optimization algorithm developed to solve multi-objective optimization problems based on the goal attainment method. This algorithm consists of transforming the original optimization problem shown on Equations 1 to 3:

$$\min_{\mathbf{x}} \mathbf{F}(\mathbf{x}) \quad (1)$$

Subject to:

$$\mathbf{g}(\mathbf{x}) \leq 0 \quad (2)$$

$$\mathbf{h}(\mathbf{x}) = 0 \quad (3)$$

into a problem given by Equations 4 to 7:

$$\min_{\gamma, \mathbf{x}} (\gamma) \quad (4)$$

Subject to:

$$\mathbf{F}(\mathbf{x}) - \gamma \cdot \mathbf{w} \leq \mathbf{F}_0 \quad (5)$$

$$\mathbf{g}(\mathbf{x}) \leq 0 \quad (6)$$

$$\mathbf{h}(\mathbf{x}) = 0 \quad (7)$$

where  $\mathbf{x}$  is defined on an open set  $\Omega \subseteq \mathfrak{R}^n$ ,  $\mathbf{F}(\mathbf{x}): \Omega \rightarrow \mathfrak{R}^m$  is a vector of objective functions,  $\mathbf{g}(\mathbf{x}): \Omega \rightarrow \mathfrak{R}^m$  defines the inequality constraints and  $\mathbf{h}(\mathbf{x}): \Omega \rightarrow \mathfrak{R}^{ne}$  the equality constraints,  $\gamma \in \mathfrak{R}$ ,  $\mathbf{w} \in \mathfrak{R}^m$  is the design weight vector, and  $\mathbf{F}_0 \in \mathfrak{R}^m$  are reference goals for the objective functions.

This modification transforms the nonlinear optimization problem into a linear objective function with nonlinear constraints. The main idea doing this transformation is to perform an optimization algorithm that looks for the solution riding over the constraints of the problem. Note that, now, the problem has an extra variable  $\gamma$ . The parameters  $\mathbf{w}$  and  $\mathbf{F}_0$  are used to get different solutions from the Pareto set.

Except in particular cases where the characteristics of the optimization problem allow a customized solution as in (Cheraghi et al. 1996), the literature indicates SLP and GRG methods to solve this kind of problem as in (Shyamsundar and Rangaiah 2000), (Han and Rangaiah 1997), (Lim and Cheraghi 1998), (Gopal and Biegler 1997) e (Syam 2005). A large amount of optimization algorithms has been developed and it will not be helpful to describe all of them here. The most important algorithms related to this work are: Successive Linear Programming (SLP) and Generalized Reduced Gradient (GRG). Due to the notoriety of these methods they will not be detailed here, but can be found in (Nocedal and Wright 1999).

## 2 Concepts of the Algorithm

The first step in the algorithm is to find where the initial point, given by the augmented vector of variables  $[x; \gamma]$ , satisfies the constraints shown in Equation 5. Since the values of  $x$  are specified, the value of  $\gamma$  is given by Equation 8:

$$\gamma = \max\left(\frac{\mathbf{F}_0 + \mathbf{F}(\mathbf{x})}{\mathbf{w}}\right) \quad (8)$$

Note that in the previous equation the maximum value is required, since this algorithm allows the optimization of multiple objective function and the maximum value guarantees that the algorithm is beginning at the right point (on the higher objective function or, in other words, at the most internal new constraint).

Once the initial value of  $\gamma$  is determined, the algorithm can be taken to the first iteration which consists of a sequence of steps for determining the best search direction, as shown in the following subsections.

### 2.1 Search Direction Determination

The first step of the proposed algorithm determines the best search direction. It is important to assume that at least one constraint is always active, considering that the set of constraints is composed by the original set and the objective functions of the previous problem, but it is possible that more than one constraint is active at same point. In this case, the best search direction to be followed must respect the active set of constraints. Further considerations will show how to inactivate unnecessary constrains. Considering that the gradient of a constraint is always normal to its surface or hyper-surface, the search direction must be orthogonal to the gradients of every active constraint.

As a second criterion, the best search direction must be as much aligned as possible to the gradient of the objective function shown in Equation 4. Indeed, a secondary optimization problem is generated here from the main optimization problem, which is easier to settle and has an analytical solution. Considering a set of active constraints, including equality and inequality constraints, given by:

$$\mathbf{R}(\mathbf{x}) = 0 \quad (9)$$

The necessary condition is the components of the search direction being orthogonal to every vector normal to the constraints or, in other words, the dot product of the Jacobian matrix of  $R(x)$  and the search direction must be a null vector (the search direction belongs to the null space of the Jacobian matrix):

$$\begin{bmatrix} \frac{\partial R_1}{\partial x_1} & \frac{\partial R_1}{\partial x_2} & \dots & \frac{\partial R_1}{\partial x_n} & \frac{\partial R_1}{\partial \gamma} \\ \frac{\partial R_2}{\partial x_1} & \frac{\partial R_2}{\partial x_2} & \dots & \frac{\partial R_2}{\partial x_n} & \frac{\partial R_2}{\partial \gamma} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial R_c}{\partial x_1} & \frac{\partial R_c}{\partial x_2} & \dots & \frac{\partial R_c}{\partial x_n} & \frac{\partial R_c}{\partial \gamma} \end{bmatrix} \begin{bmatrix} lx_1 \\ lx_2 \\ \dots \\ lx_n \\ l\gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (10)$$

Performing a Gaussian elimination in (10) in order to obtain Equation 11, it is possible to identify two vectors: the first one is a vector of dependent variables  $[l\gamma; \mathbf{lx}_D]$  and the second one is the vector of independent variables  $\mathbf{lx}_I$ .

$$\begin{bmatrix} B & 0 & 0 & \dots & 0 \\ 0 & C_1 & 0 & \dots & 0 \\ 0 & 0 & C_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & C_{nc-1} \end{bmatrix} \begin{bmatrix} l\gamma \\ lx_{D1} \\ lx_{D2} \\ \dots \\ lx_{D_{nc-1}} \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ A_{c1} & A_{c2} & A_{c3} & \dots & A_{cn} \end{bmatrix} \begin{bmatrix} lx_{I1} \\ lx_{I2} \\ lx_{I3} \\ \dots \\ lx_{I_{n-nc+1}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (11)$$

This partition occurs because, for each active constraint, one degree of freedom is lost, and a vector of  $nc$  dependent variables is determined. It is necessary that the values of the coefficients  $C$  be as different of zero as possible and, for this reason, a complete pivoting is needed when performing the Gaussian elimination. Note that the variable  $l\gamma$  is always placed in the first position of the dependent vector and the first constraint is the one resulting from the objective function of the previous problem (Equation 5).

In this manner, the secondary optimization consists of looking for the search direction vector most aligned with the gradient of the objective function of Equation 4 and orthogonal to the gradients of the constraints. As the gradient of the objective function of Equation 4 is the vector  $[0, \dots, 0, 1]$ , the objective function proposed to this secondary problem is:

$$\min_{\mathbf{lx}_I} (l\gamma) \quad (12)$$

$$\mathbf{lx}_I^T \mathbf{lx}_I + \mathbf{lx}_D^T \mathbf{lx}_D + l\gamma^2 - 1 = 0 \quad (13)$$

The constraint given in Equation 13 is necessary to guarantee that the Euclidian norm of the search direction vector is 1 and the dot product with the objective function gradient is a function of the angle between this gradient and the search direction vector only, and do not depends on the length of the search direction vector.

Rewriting the problem in the dual space, the Lagrangian function is given as follows:

$$L(\mathbf{lx}_I) = l\gamma + \lambda (\mathbf{lx}_I^T \mathbf{lx}_I + \mathbf{lx}_D^T \mathbf{lx}_D + l\gamma^2 - 1) \quad (14)$$

Using Equation (11) to write Equation (14) as a function of  $\mathbf{lx}_I$ , in matrix notation, it becomes:

$$L(\mathbf{lx}_I) = -\frac{\mathbf{A}_{1,1:\text{end}} \mathbf{lx}_I}{B} + \lambda \left( \mathbf{lx}_I^T \mathbf{lx}_I + \mathbf{lx}_I^T \mathbf{A}^T \begin{bmatrix} B & 0 \\ 0 & \mathbf{C} \end{bmatrix}^{-1T} \begin{bmatrix} B & 0 \\ 0 & \mathbf{C} \end{bmatrix}^{-1} \mathbf{A} \mathbf{lx}_I - 1 \right) \quad (15)$$

where  $\mathbf{A}_{1,1:\text{end}}$  is the first row of the matrix  $\mathbf{A}$ .

The derivative of the Lagrangian with respect to  $\mathbf{lx}_I$  is given by Equation 16, or in a more compact form by Equation 17.

$$\frac{\partial L}{\partial \mathbf{lx}_I} = -\frac{\mathbf{A}_{1,1:\text{end}}^T}{B} + \lambda \left( 2 \mathbf{I} + 2 \mathbf{A}^T \begin{bmatrix} B & 0 \\ 0 & \mathbf{C} \end{bmatrix}^{-1T} \begin{bmatrix} B & 0 \\ 0 & \mathbf{C} \end{bmatrix}^{-1} \mathbf{A} \right) \mathbf{lx}_I \quad (16)$$

$$\frac{\partial L}{\partial \mathbf{lx}_I} = -\frac{\mathbf{A}_{1,1:\text{end}}^T}{B} + \lambda \mathbf{H} \mathbf{lx}_I \quad (17)$$

Similarly, the derivative with respect to  $\lambda$  is given by Equation 18 or Equation 19.

$$\frac{\partial L}{\partial \lambda} = \mathbf{lx}_I^T \mathbf{lx}_I + \mathbf{lx}_I^T \mathbf{A}^T \begin{bmatrix} B & 0 \\ 0 & \mathbf{C} \end{bmatrix}^{-1T} \begin{bmatrix} B & 0 \\ 0 & \mathbf{C} \end{bmatrix}^{-1} \mathbf{A} \mathbf{lx}_I - 1 \quad (18)$$

$$\frac{\partial L}{\partial \lambda} = \mathbf{lx}_I^T \frac{\mathbf{H}}{2} \mathbf{lx}_I - 1 \quad (19)$$

Turning the gradient of the Lagrangian equal to zero and isolating the component  $\mathbf{lx}_I$  in Equation 17, results in:

$$\mathbf{lx}_I = \frac{\mathbf{H}^{-1}}{\lambda} \left( \frac{\mathbf{A}_{1,1:\text{end}}^T}{B} \right) \quad (20)$$

And from Equation 19:

$$\mathbf{l}\mathbf{x}_I^T \frac{\mathbf{H}}{2} \mathbf{l}\mathbf{x}_I = 1 \quad (21)$$

Substituting Equation 20 into 21:

$$\left( \frac{\mathbf{A}_{1,1:\text{end}}^T}{B} \right)^T \left( \frac{\mathbf{H}^{-1}}{\lambda} \right)^T \frac{\mathbf{H}}{2} \frac{\mathbf{H}^{-1}}{\lambda} \left( \frac{\mathbf{A}_{1,1:\text{end}}^T}{B} \right) = 1 \quad (22)$$

It is possible to solve for  $\lambda$  given:

$$\lambda = \sqrt{\frac{\mathbf{A}_{1,1:\text{end}} \mathbf{H}^{-1T} \mathbf{A}_{1,1:\text{end}}^T}{2 B^2}} \quad (23)$$

Substituting Equation 23 into Equation 20, a final search direction is obtained:

$$l\mathbf{x}_I = \frac{\text{sign}(B) \mathbf{H}^{-1} \mathbf{A}_{1,1:\text{end}}^T}{\sqrt{\frac{\mathbf{A}_{1,1:\text{end}} \mathbf{H}^{-1T} \mathbf{A}_{1,1:\text{end}}^T}{2}}} \quad (24)$$

Once the independent components  $\mathbf{l}\mathbf{x}_I$  of the search direction have been determined, it is possible to follow in this direction looking for the optimum.

## 2.2 Moving the estimation on the constraints

It is necessary to define a new problem to perform the search on the constraints, once that this is an active set method. Although the  $\mathbf{l}\mathbf{x}_I$  direction is well determined, the other components  $\mathbf{l}\mathbf{x}_D$  and  $l\gamma$  of the search vector cannot be used because the correct point  $[\mathbf{x}; \gamma]$  in the space depends on the active set of constraints, which can be non-linear. In other words, it means that performing a progression in the independent variables  $\mathbf{x}_I$ , the active set of constraints will determine the values of the dependent variables  $\mathbf{x}_D$  and  $\gamma$ . Note that for each active constraint the problem has one dependent variable. It is possible to ride on the constraints of the problem in the direction  $\mathbf{l}\mathbf{x}_I$  solving the problem composed by the active set of constraints varying the dependent variable  $\mathbf{x}_D$  and having as a parameter a coefficient  $\alpha$  that multiply the independent fraction of the search direction  $\mathbf{l}\mathbf{x}_I$ . For each value of the coefficient  $\alpha$ , which represents a step toward the solution of the optimization problem, there is a new solution of the algebraic system, which guarantees that the problem is progressing on the active constraints.

The parameter  $\alpha$  is the degree of freedom that allows the algorithm to search a better point in the optimization problem. Since the search direction vector points to the decreasing value of the objective function, the progression of this factor must be kept until the minimum of the problem be reached or a new constraint has become active.

Defining a criterion to this search, the assumption that has to be done is the minimal value of  $\gamma$  to be reached varying  $\alpha$ . In this way, considering:

$$\gamma = \frac{\mathbf{F}(\mathbf{x}) - \mathbf{F}_0}{\mathbf{w}} \quad (25)$$

The derivative of  $\gamma$  with respect to  $x$  is given by:

$$\frac{\partial \gamma}{\partial \mathbf{x}} = \frac{1}{\mathbf{w}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \quad (26)$$

And the derivative with respect to  $\alpha$  can be written by:

$$\frac{d\gamma}{d\alpha} = \frac{1}{\mathbf{w}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\alpha} \quad (27)$$

Since the vector  $\mathbf{x}$  is composed by  $\mathbf{x}_I$  and  $\mathbf{x}_D$ , being the last one dependent on  $\mathbf{x}_I$ , the expression can be divided in two terms:

$$\frac{d\gamma}{d\alpha} = \frac{1}{\mathbf{w}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}_I} \frac{d\mathbf{x}_I}{d\alpha} + \frac{1}{\mathbf{w}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}_D} \frac{\partial \mathbf{x}_D}{\partial \mathbf{x}_I} \frac{d\mathbf{x}_I}{d\alpha} \quad (28)$$

Considering that the progression in the variables  $\mathbf{x}_I$  is given by the expression:

$$\mathbf{x}_I = \mathbf{x}_{I0} + \alpha \cdot \mathbf{l}\mathbf{x}_I \quad (29)$$

The derivative of  $\mathbf{x}_I$  with respect to  $\alpha$  is:

$$\frac{d\mathbf{x}_I}{d\alpha} = \mathbf{l}\mathbf{x}_I \quad (30)$$

Since the solution rides always on the active constraints when  $\mathbf{x}_I$  is changed, the value of expression  $R(x)$  is always equal to zero to any variation of  $\mathbf{x}_I$ , and then:

$$d\mathbf{R} = 0 \quad (31)$$

As  $\mathbf{R}$  is a function of  $\mathbf{x}_I$  and  $\mathbf{x}_D$ , the Equation 31 can be written as:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{x}_I} + \frac{\partial \mathbf{R}}{\partial \mathbf{x}_D} \frac{\partial \mathbf{x}_D}{\partial \mathbf{x}_I} = 0 \quad (32)$$

Through this relation it is possible to find the derivative of  $\mathbf{x}_D$  with respect to  $\mathbf{x}_I$ :

$$\frac{\partial \mathbf{x}_D}{\partial \mathbf{x}_I} = - \left( \frac{\partial \mathbf{R}}{\partial \mathbf{x}_D} \right)^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{x}_I} \quad (33)$$

The final expression of the derivative of  $\gamma$  with respect to  $\alpha$  is given by:

$$\frac{d\gamma}{d\alpha} = \frac{1}{\mathbf{w}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}_I} \mathbf{l}_{\mathbf{x}_I} - \frac{1}{\mathbf{w}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}_D} \left( \frac{\partial \mathbf{R}}{\partial \mathbf{x}_D} \right)^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{x}_I} \mathbf{l}_{\mathbf{x}_I} \quad (34)$$

The final system to be solved is:

$$\mathbf{R}(\mathbf{x}_{I_0} + \alpha \mathbf{l}_{\mathbf{x}_I}, \mathbf{x}_D, \gamma) = 0$$

$$\frac{1}{\mathbf{w}} \left[ \frac{\partial \mathbf{F}}{\partial \mathbf{x}_I} \Big|_{(\mathbf{x}_{I_0} + \alpha \mathbf{l}_{\mathbf{x}_I}, \mathbf{x}_D, \gamma)} - \frac{\partial \mathbf{F}}{\partial \mathbf{x}_D} \Big|_{(\mathbf{x}_{I_0} + \alpha \mathbf{l}_{\mathbf{x}_I}, \mathbf{x}_D, \gamma)} \left( \frac{\partial \mathbf{R}}{\partial \mathbf{x}_D} \Big|_{(\mathbf{x}_{I_0} + \alpha \mathbf{l}_{\mathbf{x}_I}, \mathbf{x}_D, \gamma)} \right)^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{x}_I} \Big|_{(\mathbf{x}_{I_0} + \alpha \mathbf{l}_{\mathbf{x}_I}, \mathbf{x}_D, \gamma)} \right] \mathbf{l}_{\mathbf{x}_I} = 0 \quad (35)$$

This system of algebraic equations must be solved as a constrained system, considering all set of inactive constraints in the solution. In other words, it means that the solver must progress while no other constraint become active and stop at the current point if this happens. For each new active constraint the same sequence of steps must be followed, starting from the choice of the active constraint.

### 2.3 Activating and inactivating constraints

An additional calculation must be done to avoid the activation of unnecessary constraints. It is possible that a constraint be already described by another one, in a certain point, and the assumption of this described constraint can worsen the search direction. For this reason, it is necessary to evaluate, at each iteration, which constraint must become inactive.

Due to this assumption, an initial analysis of the possible search directions must be done. For each inequality constraint active on the current point, a search direction vector is calculated considering only the activated inequality constraint and the equality constraints. A search direction vector will be obtained for each activated inequality constraint. For each search direction obtained the following evaluation must be done:

$$\mathbf{Con} = \mathbf{R} \cdot \mathbf{l}_{\mathbf{x}} \quad (36)$$

If every value of  $Con$  is less than or equal to zero the search direction does not violate the other inequality constraints and all of them must be removed from the set of active constraint. If some value of  $Con$  is greater than zero, the search direction vector violates this constraint, which must be considered in the active set of constraints. If some constraint is violated, a new search direction vector is calculated considering the new constraints and the same evaluation is done until the calculated search vector does not violate any new constraints.

### 3 Solving the Algebraic Equations System (Linesearch Step)

Once the search direction is well determined, the algorithm must run toward the solution through a linesearch step. In this case, the parameter  $\alpha$  must be increased until the progression is not possible anymore. During the progression of the parameter  $\alpha$ , the system of algebraic equations composed by the active set of constraints must be settled considering that the solution cannot violate the inactive constraints.

Since the solution of the system of algebraic equations must be constrained, a particular solving method must be used. The strategy used here was to associate a conventional numerical method with some few changes to find the constrained solution of this problem. There is a particular characteristic of this problem that can be explored to perform the solution of this constrained system. If there is a constraint that is violated at the solution, it is necessary to perform a new iteration and the partial solution do not need to take the residuals values of the system to zero. The necessary condition to perform a new iteration is that the residuals of the equations  $\mathbf{R}(\mathbf{x}) = \mathbf{res}$  must be zero, but the last equation ( $dy/d\alpha$ ) may not be zero.

Considering the previous explained condition, the strategy to solve the system of algebraic equations is, first of all, to use a Newton method. Four situations can occur:

- The system has a solution that does not violate any other constraint;
- The system has a solution that violates one or more inactive constraints;
- The last equation does not have root and, consequently, the system does not have solution;
- One or more than one equation does not have root and, consequently, the system does not have solution.

The first situation is that one in which the optimization problem finishes and the solution of the algebraic system can be converted in the solution of the optimization problem. In the other three situations, the algorithm needs to continue with an iterative process and a new estimate of the solution must be provided.

In the second situation the solution of the system must be taken at the constraint, although this solution does not take the residuals of the system to zero. This situation will occur when there is a minimum in the chosen search direction that violates some inactive inequality constraint. In this case, it is necessary to bring the solution back to the constraint using a kind of bisection method. Using the both points, the initial point and the obtained solution, a sequence of divisions is done in the element  $\alpha$  in order to find the first violated constraint. The algorithm in this situation works in the following way:

- 1)  $x_1=[\gamma_0; x_{D0}; \alpha_0]; x_2=[\gamma_f; x_{Df}; \alpha_f];$
- 2)  $x_{ap}=(x_1+x_2)/2;$
- 3)  $x_{mean}$ =solution of the reduced problem eliminating the last equation and keeping  $\alpha$  constant ( $=\alpha_{ap}$ );
- 4) Evaluate the remaining constraints at the point  $x_{mean}$ ;
- 5) If the remaining constraints are violated then  $x_2=x_{mean}$ ; else  $x_1=x_{mean}$ ;
- 6) If  $\|x_2 - x_1\| < \text{tolerance}$  then  $x_{mean}$  is the solution; else go back to step 2.

The third situation will occur when there is not a minimum of the objective function, neither outside of the constraints. This situation commonly happens when the objective function is linear, or it monotonically decreases or increases, and it does not present a zero derivative in the search direction. Therefore, the system has no solution because the last equation has no root and the strategy to find the point on the constraint cannot use the solution of the algebraic system, as in situation two. Once the final point cannot be found the adopted strategy is to make a progressive linesearch in the variable  $\alpha$  beginning at the initial point  $[\gamma_0; x_{D0}; \alpha_0]$  and going toward the direction that decreases the derivative  $d\gamma/d\alpha$ . The progression algorithm consists of the following steps:

- 1)  $x_1 = [\gamma_0; x_{D0}; \alpha_0];$
- 2) If it is the first iteration  $x_{exp} = x_1 - d\gamma/d\alpha$ ; else  $x_{exp} = x_1 + 2(x_2 - x_1);$

- 3)  $x_2$  = solution of the reduced problem eliminating the last equation and keeping  $\alpha$  constant ( $=\alpha_{exp}$ );
- 4) Evaluate the remaining constraints at the point  $x_2$ ;
- 5) If the remaining constraints are not violated go back to step 2;
- 6)  $x_{ap} = (x_1 + x_2)/2$ ;
- 7)  $x_{mean}$  = solution of the reduced problem eliminating the last equation and keeping  $\alpha$  constant ( $=\alpha_{ap}$ );
- 8) Evaluate the remaining constraints at the point  $x_{mean}$ ;
- 9) If the remaining constraints are violated then  $x_2 = x_{mean}$ ; else  $x_1 = x_{mean}$ ;
- 10) If  $\|x_2 - x_1\| < \text{tolerance}$  then  $x_{mean}$  is the solution; else go back to step 6.

The fourth situation will occur when some constraint loses the sensitivity to an independent variable. Therefore, although the variation of the parameter  $\alpha$  occurs, there is no value of the dependent variables which satisfies the constraint, and it is necessary to stop the current iteration and recalculate the set of independent variables in a new iteration. Although the problem loses sensitivity using the current set of independent variables, it is necessary to estimate a new point over the constraints where, besides such point, some constraints do not present sensitivity. This determination can be done using a linesearch algorithm similar to those shown for the second and third situations, using a different stop criterion, which evaluates if the solution of the reduced problem (without the last equation and the variable  $\alpha$ ) was obtained. When this solution is not obtained the algorithm must proceed to the next iteration, evaluating new search direction and new set of independent variables using the previous solution of the reduced problem that is still on the constraints.

Considering there must be only one sequence of steps, capable to work on all situations shown previously, the final algorithm consists of, first of all, performing an expansion as shown in situation 3, followed by a compression as shown in situation 2. The condition to stop the expansion is any of this:

- The equation  $(dy/d\alpha)$  changes its signal showing that the minimum has already been achieved;
- An inactive constraint has become active; or

- The reduced system has no solution.

The same conditions are used in the compression stage to evaluate if  $x_{mean}$  must either substitute  $x_1$  or  $x_2$ .

## 4 Comparison with Other Algorithms

This algorithm presents some similarities with other ones available, like the Generalized Reduced Gradients (GRG). However some conceptual differences are important to be highlighted. The first difference occurs in the selection of the set of independent variables. The GRG algorithm does not indicate a clear way to select the set of independent variables. In this algorithm a pivoting strategy leads to the best set of independent variables, once the highest value chosen as a pivot indicates the less sensible variable to each constraint which must be considered as an independent variable.

The second difference is the choice of the search direction vector. This algorithm uses the search direction vector shown in Equation 24, while the search direction vector used by the GRG method is given by:

$$\mathbf{l}_{\mathbf{x}_I} = -\frac{d\mathbf{F}}{d\mathbf{x}_I} + \left(\frac{d\mathbf{R}}{d\mathbf{x}_I}\right)^T \left(\frac{d\mathbf{R}}{d\mathbf{x}_D}\right)^{-1T} \frac{d\mathbf{F}}{d\mathbf{x}_D} \quad (37)$$

In order to illustrate the difference between these two ways to obtain the search direction, consider the following example:

$$F(\mathbf{x}) = 3x_1 + 4x_3 \quad (38)$$

s.t.:

$$x_1^2 + (x_2 + 3)^2 + x_3^2 - 25 \leq 0 \quad (39)$$

The minimum of this problem occurs at the point  $x^* = [-3, -3, -4]$  with  $F(x^*) = -25$ . Considering an initial guess  $x_0 = [-2.4, 0, -3.2]$ , the determination of the search direction obtained by the proposed algorithm begins obtaining the gradient of the modified problem, as follows:

$$\frac{\partial R}{\partial x} = \begin{bmatrix} 3 & 0 & 4 & 1 \\ -4.8 & 6 & -6.4 & 0 \end{bmatrix} \quad (40)$$

Doing the complete pivoting, the system becomes:

$$\begin{bmatrix} 1 & 0 \\ 0 & -6.4 \end{bmatrix} \begin{bmatrix} \mathbf{I}_y \\ \mathbf{I}_x \end{bmatrix} + \begin{bmatrix} 3.75 & 0 \\ 6 & -4.8 \end{bmatrix} \begin{bmatrix} \mathbf{I}_x \\ \mathbf{I}_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (41)$$

which leads to the  $H$  matrix given by:

$$\begin{aligned} \mathbf{H} &= 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 2 \begin{bmatrix} 3.75 & 0 \\ 6 & -4.8 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & -6.4 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -6.4 \end{bmatrix}^{-1} \begin{bmatrix} 3.75 & 0 \\ 6 & -4.8 \end{bmatrix} \\ &= \begin{bmatrix} 31.8828 & -1.4062 \\ -1.4062 & 3.1250 \end{bmatrix} \end{aligned} \quad (42)$$

And finally leads to the search direction:

$$\mathbf{I}_x = \frac{\begin{bmatrix} 31.8828 & -1.4062 \\ -1.4062 & 3.1250 \end{bmatrix}^{-1} \begin{bmatrix} 3.75 \\ 0 \end{bmatrix}}{\sqrt{\frac{\begin{bmatrix} 3.75 \\ 0 \end{bmatrix}^T \begin{bmatrix} 31.8828 & -1.4062 \\ -1.4062 & 3.1250 \end{bmatrix}^{-1} \begin{bmatrix} 3.75 \\ 0 \end{bmatrix}}{2}}} = \begin{bmatrix} -0.2530 \\ -0.1138 \end{bmatrix} \quad (43)$$

In the GRG algorithm the search direction obtained, considering the same independent variables, is:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{x}_1} = \begin{bmatrix} \frac{\partial \mathbf{R}}{\partial x_2} & \frac{\partial \mathbf{R}}{\partial x_1} \end{bmatrix} = [6 \quad -4.8] \quad (44)$$

$$\frac{\partial \mathbf{R}}{\partial \mathbf{x}_D} = \frac{\partial \mathbf{R}}{\partial x_3} = -6.4 \quad (45)$$

$$\frac{\partial S}{\partial \mathbf{x}_1} = \begin{bmatrix} \frac{\partial S}{\partial x_2} & \frac{\partial S}{\partial x_1} \end{bmatrix} = [0 \quad 3] \quad (46)$$

$$\frac{\partial S}{\partial x_D} = \frac{\partial S}{\partial x_3} = 4 \quad (47)$$

$$\mathbf{I}_x = \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 6 \\ -4.8 \end{bmatrix} \cdot (-6.4)^{-1} \cdot 4 = \begin{bmatrix} -3.75 \\ 0 \end{bmatrix} \quad (48)$$

Both search directions are linearly independent in this case, what means that these algorithms have different ways to obtain the search direction.

The third difference of the GA-GRG algorithm and GRG method is the transformation of the main problem in a linear problem increasing the number of constraints. Therefore it is possible to use a multiple objective optimization where each the objective functions are transformed in constraints. This transformation leads the minimum of the problem to lays always over, at least, one constraint, which takes the algorithm to use only one kind of iteration, different from others which uses different kinds of iteration depending on where is the optimum (i.e., active constraint or not).

## 5 Results

The evaluation of this algorithm was performed using a set of benchmark problems obtained from the literature and comparing the behavior of the proposed algorithm with the FMINCON routine developed in MATLAB<sup>®</sup> 5.3 Release 11. This routine consists of a SQP method based on the active set method. The algorithm GA-GRG was also developed in MATLAB environment to allow a better comparison. The following sequence of benchmark problems was used:

Problem 1 (Ryoo and Sahindis 1995):

$$F(x) = x_1 + x_2 + x_3 \quad (49)$$

Subject to:

$$100000(x_4 - 100) - 120x_1(300 - x_4) = 0 \quad (50)$$

$$100000(x_5 - x_4) - 80x_2(400 - x_5) = 0 \quad (51)$$

$$100000(500 - x_5) - 40x_3(600 - 500) = 0 \quad (52)$$

$$[0, 0, 0, 100, 100] \leq x \leq [15834, 36250, 12500, 300, 400] \quad (53)$$

Solution (approximation):  $x^* = [158.34 \ 1644.3155 \ 5394.3167 \ 131.9339 \ 284.2273]$

Problem 2 (Biegler and Cuthrell 1985):

$$F(x) = x_2 \quad (54)$$

Subject to:

$$1 + x_1 - x_2^2 \leq 0 \quad (55)$$

$$1 - x_1 - x_2^2 \leq 0 \quad (56)$$

$$\frac{1}{2} - x_2 \leq 0 \quad (57)$$

Solution:  $x^* = [0 \ 1]$

Problem 3 (Edgar and Himmelblau 1989):

$$F(x) = x_1 x_2 \quad (58)$$

Subject to:

$$-x_1 - x_2 \leq 0 \quad (59)$$

$$x_1^2 + x_2^2 - 25 \leq 0 \quad (60)$$

Solution:  $x^* = [-\sqrt{12.5} \ , \ \sqrt{12.5}]$

Problem 4 (Ryoo and Sahindis 1995):

$$F(x) = x_1 + x_2 \quad (61)$$

Subject to:

$$x_1^2 + x_2^2 - 4 \leq 0 \quad (62)$$

$$-x_1^2 - x_2^2 + 1 \leq 0 \quad (63)$$

$$x_1 - x_2 = 0 \quad (64)$$

$$\text{Solution: } x^* = [-\sqrt{2}, -\sqrt{2}]$$

Problem 5 (Ryoo and Sahindis 1995):

$$F(x) = -x_1 - x_2 \quad (65)$$

Subject to:

$$x_1 x_2^2 - 4 \leq 0 \quad (66)$$

$$x_1 - 6 \leq 0 \quad (67)$$

$$x_2 - 4 \leq 0 \quad (68)$$

$$-x_1 \leq 0$$

$$-x_2 \leq 0 \quad (69)$$

$$\text{Solution: } x^* = [16, 4/6];$$

Problem 6 (Ryoo and Sahindis 1995):

$$F(x) = x_1^{0.6} + x_2^{0.6} - 6x_1 - 4x_3 + 3x_4 \quad (70)$$

Subject to:

$$x_1 + 2x_3 - 4 \leq 0 \quad (71)$$

$$x_2 + 2x_4 - 4 \leq 0 \quad (72)$$

$$x_1 - 3 \leq 0 \quad (73)$$

$$x_2 - 4 \leq 0 \quad (74)$$

$$x_3 - 2 \leq 0 \quad (75)$$

$$x_4 - 1 \leq 0 \quad (76)$$

$$-x_1 \leq 0 \quad (77)$$

$$-x_2 \leq 0 \quad (78)$$

$$-x_3 \leq 0 \quad (79)$$

$$-x_4 \leq 0 \quad (80)$$

$$-3x_1 + x_2 - 3x_3 = 0 \quad (81)$$

Solution:  $x^* = [0, 0, 0, 0]$ ;

Problem 7 (Ryoo and Sahindis 1995):

$$F(x) = 29.4x_1 + 18x_2 \quad (82)$$

Subject to:

$$0.2458x_1^2 - x_1x_2 + 6x_2 \leq 0 \quad (83)$$

$$x_1 - 15.8 \leq 0 \quad (84)$$

$$x_2 - 30 \leq 0 \quad (85)$$

$$-x_1 \leq 0$$

$$-x_2 \leq 0$$

Solution (approximation):  $x^* = [8.18, 7.56]$ ;

The evaluation of the algorithm was performed through the following criteria:

### 5.1 Effectiveness

The effectiveness represents the relation among the time spent to solve the set of problems. The most effective algorithm will be that one which spends less time to find the solution, determined by:

$$Ez_j = \frac{100}{N} \sum_{i=1}^N \frac{\min_k T_{i,k}}{T_{i,j}} \quad (86)$$

where  $Ez_j$  is the effectiveness of the algorithm  $j$ ,  $N$  is the number of problems evaluated,  $T_{i,j}$  is the time spent to solve the problem  $i$  by the algorithm  $j$ .

## 5.2 Efficiency

The efficiency represents the behavior of the algorithm related to the number of evaluations of the objective function. The efficiency will be as better as less evaluations of the objective function are done.

$$Ef_j = \frac{100}{N} \sum_{i=1}^N \frac{\min_k NS_{i,k}}{NS_{i,j}} \quad (87)$$

where  $Ef_j$  is the efficiency of the algorithm  $j$ ,  $N$  is the number of problems evaluated,  $NS_{i,j}$  is the number of objective function evaluations in the problem  $i$  by the algorithm  $j$ .

## 5.3 Robustness

The robustness is a parameter associated with the capacity of the algorithm to solve different problems. This parameter can be expressed by:

$$Rb_j = \frac{100Ni_j}{N} \quad (88)$$

where  $Rb_j$  is the robustness of the algorithm  $j$ ,  $N$  is the number of problems evaluated,  $Ni_j$  is the number of problems effectively solved by the algorithm  $j$ .

## 5.4 Quality

Quality is a parameter related to the approximation of the solution to the best solution found to one problem. It is expressed by:

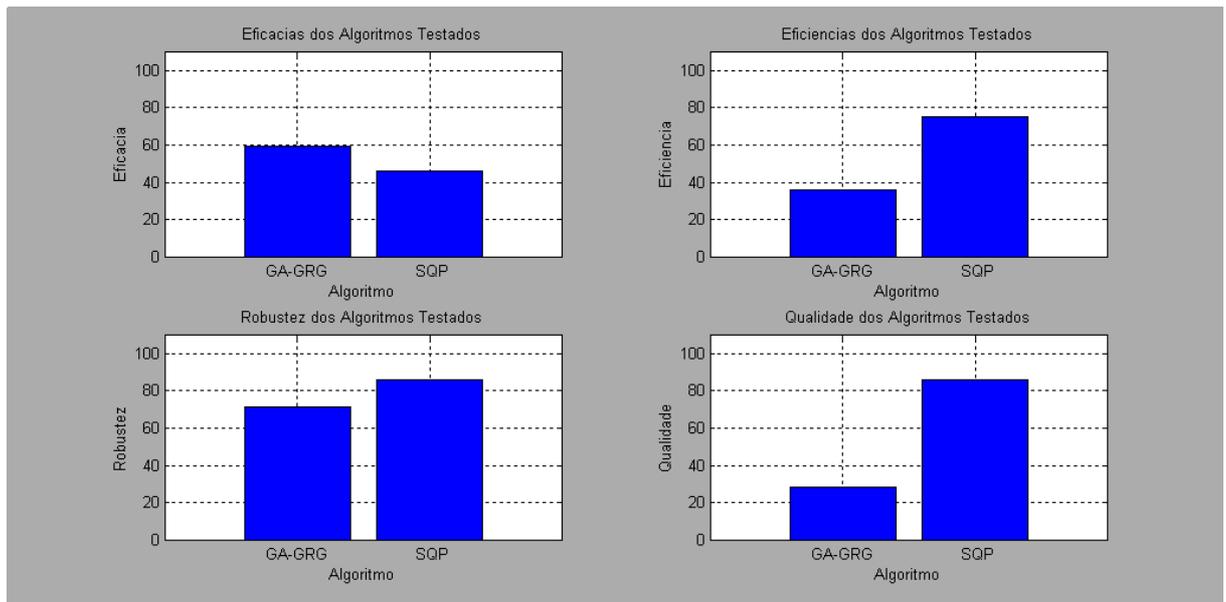
$$Ql_j = \frac{100}{N} \sum_{i=1}^N \frac{\min_k DS_{i,k}}{DS_{i,j}} \quad (89)$$

where  $Ql_j$  is the quality of the algorithm  $j$ ,  $N$  is the number of problems evaluated,  $DS_{i,j}$  is the distance between the optimal solution of the problem  $i$  and the result obtained by the algorithm  $j$ , defined as:

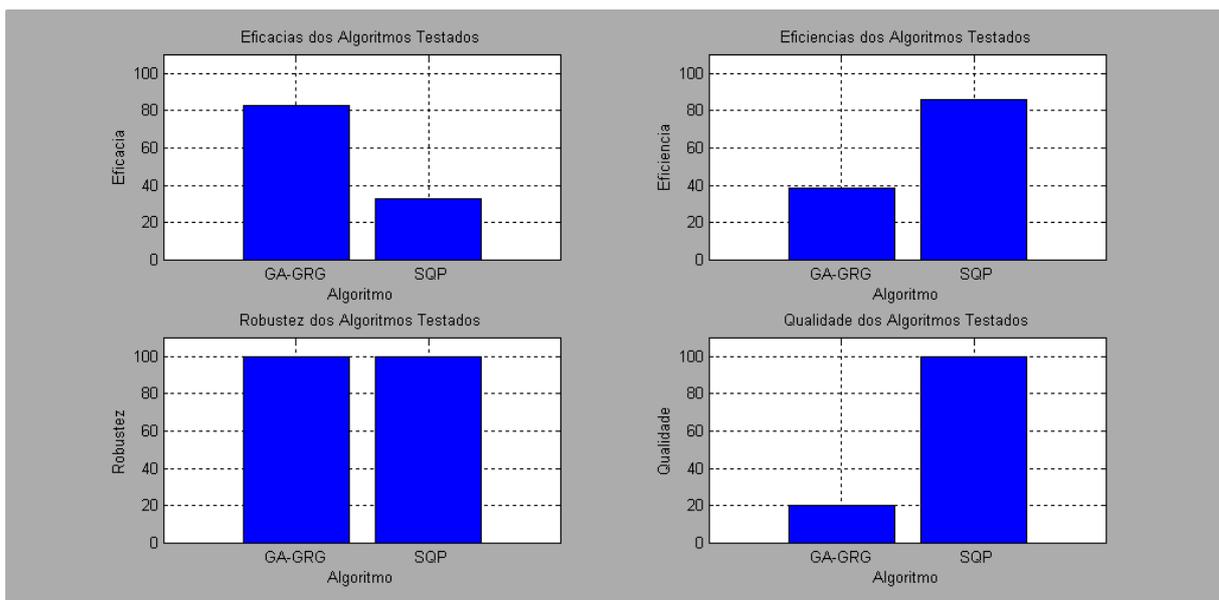
$$DS_{i,j} = \frac{\sqrt{(\mathbf{x}_{i,j}^* - \mathbf{x}_i^{ot})^T (\mathbf{x}_{i,j}^* - \mathbf{x}_i^{ot})}}{TolX_i} + \frac{\sqrt{(S_{i,j}^* - S_i^{ot})^T (S_{i,j}^* - S_i^{ot})}}{TolS_i} \quad (90)$$

where  $\mathbf{x}_{i,j}^*$  is the solution obtained using algorithm  $j$  for the problem  $i$ ;  $S_{i,j}^*$  is the best value for the objective function obtained using algorithm  $j$  for the problem  $i$ ;  $\mathbf{x}_i^{\text{ot}}$  is the best solution for the problem  $i$ , which, in general, is the algebraic solution;  $S_i^{\text{ot}}$  is the value of the objective function for the solution  $\mathbf{x}_i^{\text{ot}}$ ; TolX $_i$  is the tolerance for the solution  $\mathbf{x}$  and TolS $_i$  is the tolerance for the value of the objective function.

Figure 1 summarizes the results obtained to the analysis of the four parameters.



**Figure 1:** Evaluation of the Efficiency, Effectiveness, Quality, and Robustness of the Algorithm GA-GRG (1) and the algorithm SQP (2), using problems from 1 to 7.



**Figure 2:** Evaluation of the Efficiency, Effectiveness, Quality, and Robustness of the Algorithm GA-GRG (1) and the algorithm SQP (2), using problems from 2 to 6.

## 6 Conclusions

Although the GA-GRG algorithm presents a sequence of calculation that, at first view, seems to be more complex and seems to take a long time to be executed, the results shown that the effectiveness of this algorithm is better than the SQP algorithm. This observation shows that, in most cases, GA-GRG is faster than the SQP used in this comparison. On the other hand, the other parameters of comparison show that SQP is more robust, more efficient, and presents a better quality of the solution. In other words, it is possible to see that SQP solves more problems, makes less evaluations of the objective function, and produces a solution more precise than the GA-GRG algorithm.

Although, of the above characteristics of the algorithm GA-GRG be dissatisfactory, the most important characteristic for real-time optimization is the effectiveness, once the time spent to solve a problem is the bottleneck condition for this kind of application.

The efficiency, which is related to the number of objective function evaluations, is an important parameter when the objective function is very complex, with a lot of calculations involving the solution of dynamic equations represented by systems of differential-algebraic equations. It means that this algorithm can be well applied for real-time optimization since the model which represents the process is usually simplified.

The reduction in robustness of the algorithm GA-GRG is due to the problems 1 and 7 which could not be solved by this algorithm. This can easily be seen in Figure 2 where the same tests were performed without problems 1 and 7. The reason why these problems could not be solved is related to the presence of bilinear constraints, which limits the solution of this algorithm. Bilinear constraints involve a large change of search direction that cannot be captured by the algorithm, once that these constraints represent an asymptotic approximation of two orthogonal axes. The improvement of this algorithm should involve a modification in the line search stage to turn possible the work with bilinear constraints.

Although the quality of the solutions obtained by the GA-GRG algorithm are worst than the SQP, it is not too important for industrial applications. If the problem is correctly scaled, this parameter do not impacts too much in this analysis. Of course, it is important to mention

that the GA-GRG algorithm should not be applied for problems which need a high accuracy in the solution.

## 7 Reference

Aboelela E, and Douligeris C. (1999). Fuzzy generalized network approach for solving an optimization model for routing in B-ISDN. *Telecommunication Systems* 12, 237-263.

Biegler LT, and Cuthrell JE. (1985). Improved Unfeasible Path Optimization for Sequencia Modular Simulators II - The optimization algorithm. *Computer and Chemical Engineering* 9, 257-267.

Cheraghi SH, Lim HS, and Motavalli S. (1996). Straightness and Flatness Tolerance Evaluation: An Optimization Approach. *Precision Engineering* 18, 30 - 37.

Chiam SC, Tan KC, and Mamun AA. (2007). Multiobjective Evolutionary Neural Networks for Time Series Forecasting. *Lecture Notes in Computer Science* 4403, 346–360.

Coello CAC. (1998). Using the Min-Max Method to Solve Multiobjective Optimization Problems with Genetic Algorithms. *Lecture Notes in Artificial Intelligence* 1484, 303–313.

Day RO, and Lamont GB. (2005). Multiobjective Quadratic Assignment Problem Solved by an Explicit Building Block Search Algorithm – MOMGA-IIa. *Lecture Notes in Computer Science* 3498, 91–100.

Edgar TF, and Himmelblau DM. (1989). *Optimization of Chemical Process*. McGraw-Hill, Singapore.

Gopal V, and Biegler LT. (1997). Nonsmooth Dynamic Simulation with Linear Pragramim Based Method. *Computers and Chemical Engineering* 21(7), 675 - 689.

Han G, and Rangaiah GP. (1997). A method for Calculation vapour-liquid and liquid-liquid equilibrium. *Computers and Chemical Engineering* 21, 905 - 913.

Kleemann MP, Lamont GB, Cooney A, and Nelson TR. (2007). A Multi-tieredA Multi-tiered Memetic Multiobjective Evolutionary Algorithm for the Design of Quantum Cascade Lasers. *Lecture Notes in Computer Science* 4403, 186-200.

Kukkonen S, and Lampinen J. (2004). An Extension of Generalized Differential Evolution for Multi-objective Optimization with Constraints. *Lecture Notes in Computer Science* 3242, 752-761.

Kumar R, Singh PK, and Chakrabarti PP. (2004). Improved Quality of Solutions for Multiobjective Spanning Tree Problem Using Distributed Evolutionary Algorithm. *Lecture Notes in Computer Science* 3296, 494–503.

Lim HS, and Cheraghi SH. (1998). An Optimization Approach to Shape Matching and Recognition. *Computers and Chemical Engineering* 24, 183 - 200.

Mezura-Montes E, and Coello CAC. (2004). An Improved Diversity Mechanism for Solving Constrained Optimization Problems Using a Multimembered Evolution Strategy. *Lecture Notes in Computer Science* 3102, 700–712.

Nocedal J, and Wright S. (1999). *Numerical Optimization*. Springer Verlag.

Parks GT, and Miller I. (1998). Selective Breeding in a Multiobjective Genetic Algorithm. *Lecture Notes in Computer Science* 1498, 250-259.

Roth S, Gepperth A, and Igel C. (2006). Multi-Objective Neural Network Optimization for Visual Object Detection. *Studies in Computational Intelligence (SCI)* 16, 629-655.

Ryoo HS, and Sahindis NV. (1995). Global Optimization of Nonconvex NLPs and MINLPs with Application in Process Design. *Computer and Chemical Engineering* 19, 551-556.

Sareni B, Regnier J, and Roboam X. (2004). Recombination and Self-Adaptation in Multi-objective Genetic Algorithms. *Lecture Notes in Computer Science* 2936, 115–126.

Shyamsundar V, and Rangaiah GP. (2000). A Method for Simulation and Optimization of Multiphase Distillation. *Computers and Chemical Engineering* 24, 23 - 37.

Syam MI. (2005). Nonlinear Optimization Exclusion Tests for Finding All Solutions of Nonlinear Equations. *Applied Mathematics and Computation* 170, 4104 - 4116.

Wang JQ, J.Qin, and Kang LS. (2006). A New QoS Multicast Routing Model and Its Immune Optimization Algorithm. *Lecture Notes in Computer Science* 4159, 369-378.



## Capítulo 4

### Conversão de Modelos

Conforme mostrado no capítulo anterior verificou-se que, embora modificações tenham sido realizadas em um algoritmo de otimização com o intuito de torná-lo mais rápido na resolução de problemas de otimização, pouco ganho foi obtido neste sentido. Considerando que, com o intuito de se proporcionar otimização em tempo real na mesma camada de controle (monocamada), necessita-se de uma resposta na ordem de minutos e que otimizadores de processos que utilizam modelos fenomenológicos para a otimização trabalham na ordem de horas, conclui-se que a solução da otimização em tempo real em monocamada não está na modificação de algoritmos de otimização. O grande problema encontrado está no fato de que os otimizadores realizam inúmeras avaliações da solução do modelo que, por sua vez, torna-se extremamente morosa em se tratando de modelos fenomenológicos.

A proposta estudada neste capítulo trata da concepção de um método de conversão de modelos fenomenológicos em formulações simplificadas de rápida avaliação capazes de transformar soluções que podem levar até mesmo horas, no caso de modelos fenomenológicos, em soluções que sejam geradas em poucos segundos. A metodologia estudada apresenta um exemplo de aplicação prática e mostra os seus resultados indicando as suas potencialidades e todos os seus problemas a serem resolvidos. Os detalhes deste estudo estão no artigo apresentado no apêndice B.



## **Apêndice B**

**Artigo: “Method for Conversion of First Principle Models into Simplified Formulations for Application in Dynamic Real-Time Optimization”**



# **METHOD FOR CONVERSION OF FIRST PRINCIPLE MODELS INTO SIMPLIFIED FORMULATIONS FOR APPLICATION IN DYNAMIC REAL-TIME OPTIMIZATION**

## **1 Introduction**

Process control and optimization is strongly dependent on a previous effort that involves a wide knowledge about the plant behavior. Such knowledge cannot be acquired without a plant identification or development of a good process model. Nowadays, there are a lot of techniques to identify linear models which can give to the control designer a good idea of the process characteristics in most cases where there are no extreme nonlinearities. However, in some cases, the plant changes considerably its behavior and a linear model is not enough to describe the process on the desired operating points. In other cases, although such operating points can be described by only one linear model, two different operating points of the same process are described by two different linear models, what do not allow a transition by an optimizer. In such cases, it is necessary to use a non-linear optimization and for consequence, a nonlinear model has to be obtained.

Linear system identification has been thoroughly studied in the past two to three decades and there is a vast amount of literature dedicated to this. It is fair to say that the major questions concerning the general concepts of linear system identification such as selection of input signal, structure selection, model validation are fairly well understood and there are not many theoretical challenges in this field. And the existent questions are constrained to the industrial environment. The probable reason for the wide popularity of linear models may be due to the applicability of the principle of superposition and homogeneity to a linear system. Further, in short, no mathematics beyond the theory of vector spaces and linear operators is required for a rather complete account of linear system identification.

Unfortunately, no such easiness exists for nonlinear systems. The inter relationship between response, nonparametric model, parametric model, and

system structure are far more complex when the system is nonlinear. However, times change, sufficient progress has been made in nonlinear system identification in the recent past to give some degree of confidence in approaching several classes of nonlinear problems.

Dynamic Real-Time optimization supposes that the process is completely represented and all of its operational points can be described by a nonlinear model which is able to represent all relations among variables, inversion of the signal of the gain, change of importance of variables related to the others. The most comprehensive way to represent a nonlinear process with those characteristics is using a system of Algebraic-differential equations which can include partial differential equations.

Due to the complexity of this system and the difficulty to find an optimized solution using this kind of model, it is usual to simplify the model looking for trade-off between reproducibility of the real process and the computational cost to its solution.

The first very usual simplification, which is done in order to turn the calculation feasible, is to reduce the number of differential variables. This simplification do not takes to a significant loss of quality of the model. The optimizers usually works with models composed by ordinary differential equations where the differential variable reduces to the time and all other space variables are discretized in multiple states, as often as need.

Even though this simplification is done, first principle models describes by differential-algebraic systems are still very expensive in both senses of the word. This kind of model can take a long time to be simulated and a long time to be built and validated, i.e. distillation columns which can have hundreds of states. A first principle model can take months to be built which not always is allowed in industrial environment.

Order-reduction methods have been studied and developed to simplify dynamic models as shown in (Hedengren and Edgar 2005). This kind of procedure reduces the size of the system and, hence it reduces computational time. However the difficulty to obtain the model is still the same.

Alternatives to the first principle models can be found in the literature, as well as its nonlinear identification methodology. In (Fernandes 2006), a new method for nonlinear model identification is shown. This methodology is based on the interpolation of linear model identified on different operational points of the process. This technique consists of determining an autonomous system given by the equation:

$$\frac{dx}{dt} = A(u) \cdot (x - \Omega(u)) \quad (1)$$

In this autonomous system,  $A(u)$  is the interpolation of matrix  $A$  of the state-space for each linear model identified on each operational point, choosing the variable  $u$  as independent variable in the interpolation. The vector  $\Omega(u)$  is the nonlinear function which describes the interpolation of absolute steady-state values of the different identified linear models. The interpolation of this steady-state must be done as a function of  $u$  (inputs of the system). The vector  $x$  is the states of the model. However, this method presents some practical problems to its application. The first problem is to obtain the linear models with the same order. Considering that the interpolation of the elements of the matrix  $\Omega(u)$  is done on each term, it is necessary that all linear models have the same dimension. It means that all linear models must contain the same number of states or, in other words, all linear models must be of the same order. This requirement is not always possible to satisfy, since it is possible to describe the process through different linear models with different levels of complexity (e.g., orders, structures).

Even though all linear models are obtained with the same order, there is no guarantee that the states obtained in each model conserve a relation among them, making possible a direct interpolation of those states. For instance, it is possible to suppose a case of the model composed by states  $x_{A1}$  and  $x_{A2}$  at the operating point  $A$  and states  $x_{B1}$  and  $x_{B2}$  at operating point  $B$ . It is not possible to affirm that the best interpolation of this two models is done coupling states  $x_{A1}$  and  $x_{B1}$  into a function and  $x_{A2}$  and  $x_{B2}$  into another one. It is possible that the best interpolation function is composed by  $x_{A1}$  and  $x_{B2}$  and no technique exists to determine which is the best alternative.

A third problem that occurs in this kind of model construction is related to the way in which matrices  $A$  are interpolated. Even if the different models have

been determined with the same orders and a reasonable interpolation function have been obtained, the direct interpolation of the values of matrices  $A$  can take its eigenvalues to positive values, for specific values of  $u$  (i.e., there is no guarantee that the state  $x_1$  of the model 1 can be interpolated with state  $x_1$  of model 2. Perhaps the best interpolation is  $x_1$  of model 1 with  $x_2$  of model 2), what can produce unstable interpolated models.

In this paper, a method to convert a first principle model or an identified nonlinear model into a simplified dynamic model is proposed. Such method can be evaluated quickly, making possible its application in real time optimization. Note that it is not a reduction-order method, but a conversion method that results sometimes in a model with a bigger order than the previous model. The advantage of using this kind of conversion is the generation of a dynamic algebraic model which can be quickly evaluated.

## **2 | Dynamic models based on steady-state Characteristics**

As shown in previous section the model proposed by (Fernandes 2006) seems to be a good representation for nonlinear process, however three limitations are arises and do not allow its practical application:

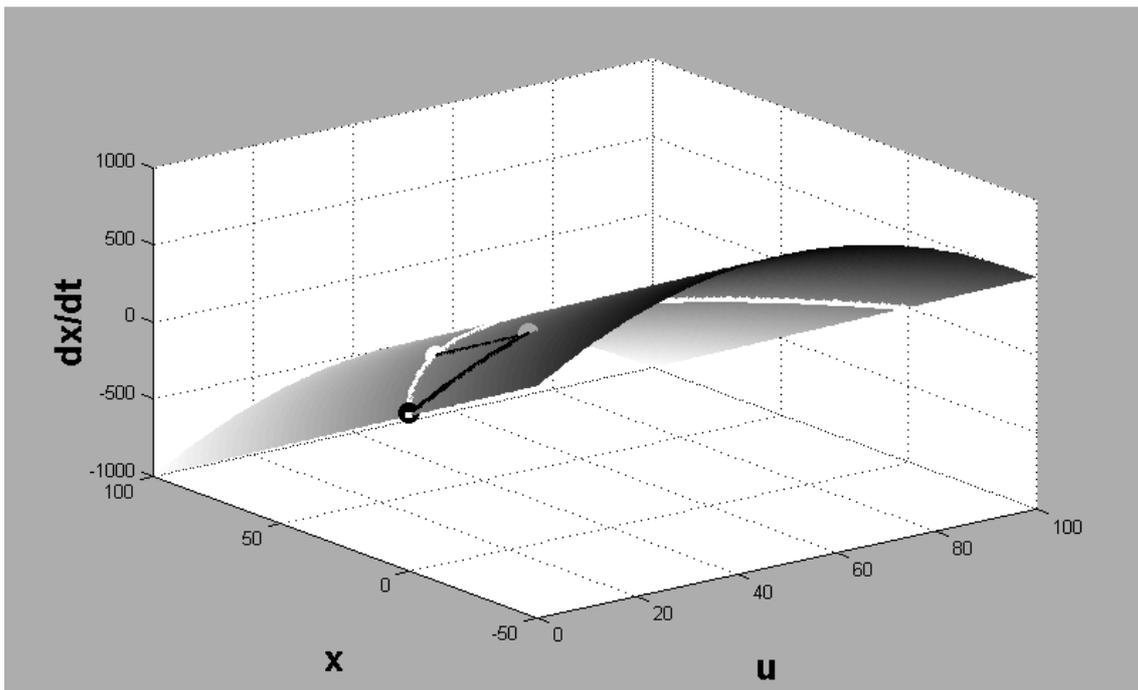
- There is no guarantee that the states in the different identified models keep a straight relation among them;
- The linear models identified must contain the same number of states
- The straight interpolation of the elements of  $A$  matrix can generate a unstable system.

The objective of this work is to reformulate the model proposed by (Fernandes 2006). The main modification is to change the original representation of the model, which is based on state space formulation where matrices are interpolated, to a monovariable representation using transfer function representation. To introduce the approach, first it will be analyzed the monovariable case and later will be discussed the generalization to the multivariable case.

The proposition of this kind of model can be understood considering its origin. Therefore, consider a monovariable system of first order whose the differential equation is given by:

$$\frac{dx}{dt} = f(x, u) \quad (2)$$

Where  $x$  is the current state of the system and, in this case, it is also the output variable; and  $u$  is the input variable. Figure .1 shows an example of a tridimensional plot of  $f(x, u)$  as a function of  $x$  and  $u$



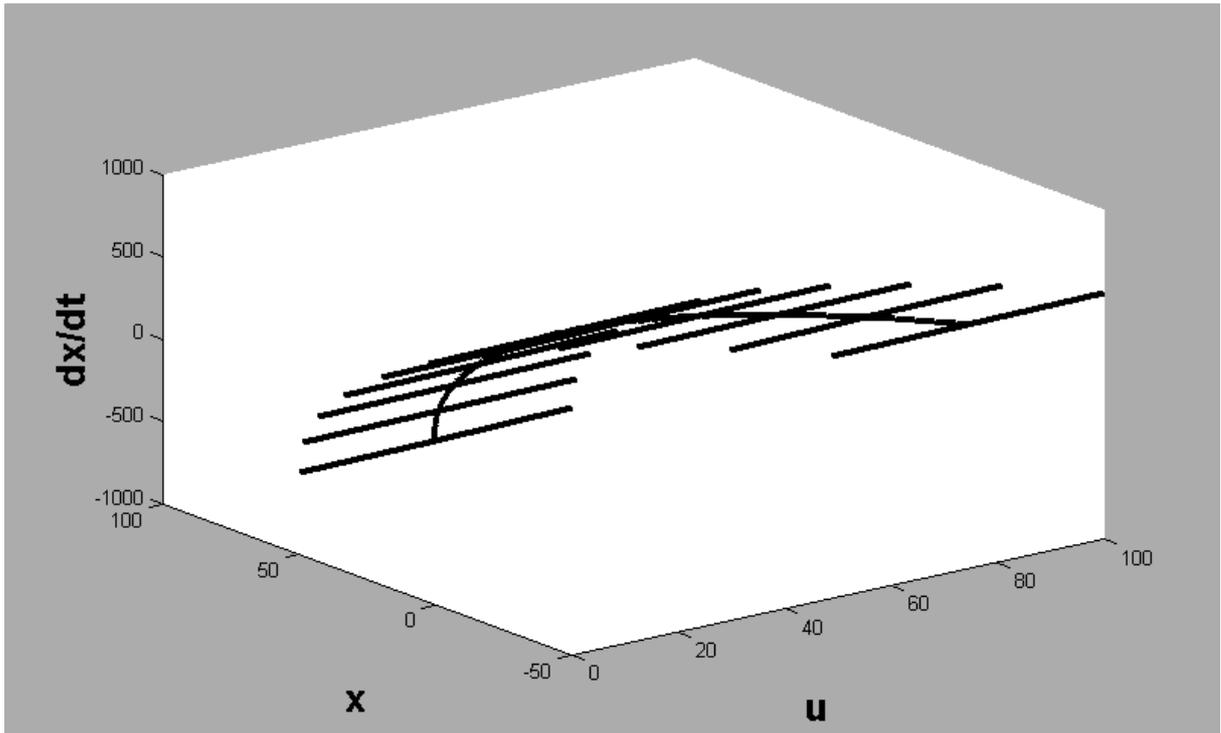
**Figure 1:** Plot of  $f(x, u)$  as a function of  $x$  and  $u$ .

Figure .1 shows a representation of the surface  $f$  (or  $dx/dt$ ) as a function of  $x$  and  $u$ . Although different systems can present different surfaces some characteristics are similar for all of them. The first consideration is that this methodology is supposed to represent stable models. Although not all industrial process are stable, such stability can be achieved by closing a control loop of the system in order to make the optimizer work on a cascade mode with this controller. Therefore, the surface must be always convergent related to variable  $x$  (negative inclination for the variable  $x$ ) on the steady-state points. In other words it is necessary that the derivative  $df/dx$  must be less than zero over the points where the term  $dx/dt$  was zero. It means that on the steady-state the poles of the system are negatives (i.e. the poles of the system is expressed by

the derivative  $df/dx$  on the steady-states  $dx/dt = f = 0$ ). Of course, for non linear systems this derivative can change for different points of the domain of the function  $f$  and the stability can only be guaranteed if in there is a steady-state where the derivative  $df/dx$  is negative for a fixed value of  $u$ .

In figure .1 the white transversal line represents the set of steady-states of the system as a function of the input  $u$ . The black and white points represent two different steady-states. A change from the operating point black to the operating point white can be represented by the black line which reaches the gray point that is not placed on the line of steady-states. When the disturbance occurs, the value of variable  $u$  is instantly changed, while variable  $x$  keeps the same value and the system reaches a transient state represented by the gray point. At the gray point the value of  $dx/dt$  is not zero and hence this term becomes the driving force to bring the system to the steady-state back again. After this disturbance the system is taken to a steady-state in a dynamic way from the gray point to the white point. In this transformation the first part of the black line between the points black and gray is an instantly transformation while the second part of the line between the points gray and white corresponds to the dynamic transformation of the system between two different operating points.

Through the observation of this transition example, it is possible to conclude that the important region to be described by the nonlinear model is around the steady-state line and the main dynamic line which takes the system to the steady-state, because the system must stand at this line or around that. It is possible to observe that the dynamic transition occurs following the dynamic of the final point represented by the line which joins the gray point to the white point. Hence it is possible to divide the representation of the system in two different parts – The steady-state line and the dynamic lines related to the different steady-state points as shown in figure 2.



**Figure 2:** Plot of  $f(x,u)$  as a function of  $x$  and  $u$ . Important parts of the function

Therefore, the determination of the model  $f$  of the process through this methodology needs a determination of a steady-state model followed by a determination of a sequence of convergent lines, called here by dynamic lines. An easy way to propose this kind of model can be represented by the following equation:

$$\frac{dx}{dt} = p(u) \cdot (x - ss(u)) \quad (3)$$

where  $ss(u)$  is the expression of the steady-state of the system and  $p(u)$  is the poles of the system to different values of inputs.

As mentioned previously, this model is assuming a system with:

- one input;
- one output;
- first order relationship;
- stable models; and
- real poles.

The following subsections will show how this representation can be generalized to become more comprehensive.

## 2.1 **Multivariable Generalization**

The consideration of a monovariable system done until here was important to understand in a graphical way the behavior of the system. However this representation can be immediately extended to the multivariable by considering the variable  $u$  as a vector of inputs and the functions  $p(u)$  and  $ss(u)$  as functions of  $\mathbb{R}^{nu} \rightarrow \mathbb{R}$ . In the same way multiple functions can be determined in order to cover all output variables. Considering that every output variable is independent of the others outputs, the identification procedure can be made as a MISO system (Multiple-Input-Single-Output) and grouping the results in a  $\mathbb{R}^{nu} \rightarrow \mathbb{R}^{ny}$  function.

Note that the proposition of the model in this way is very close to the proposed in (Fernandes 2006), however the assumption which is done here is that the matrix  $A$  assumes a diagonal shape. This assumption turns easy the interpolation of the matrix  $A$ . As shown previously, straight interpolation of elements of matrix  $A$  can generate, eventually, nonnegative eigenvalues to the interpolated matrix which is very difficult to control. In a diagonal matrix the eigenvalues appears simply on the diagonal making easier to establish constraints to the optimization routine which fits the data of the poles in a expression of interpolation.

## 2.2 **Order increasing of the model**

Systems with order bigger than one, which do not show an oscillatory behavior, can be approximated as summation or convolution of responses of first order. In general, high order appears in a dynamic system in two ways:

- Due to the addition of two different effects caused by a disturbance
- Due to sequential physical effects, where a disturbance causes an effect and the result of this effect causes a second effect which is taken as an output of the system.

The addition of effects is well explained analyzing the effect of feeding solid fuel in boilers. In general this system shows a little inverse response (RHP) and its appearance can be easily understood analyzing what really physically occurs in the process. The fuel fed in the process is always in a lower temperature than the amount of fuel in the boiler. Increasing of fuel feeding implies in an increasing in the amount of fuel mass in the boiler with less energy than the previous mass within that one. This mixture of cold and hot fuel cause a decreasing of the temperature of the combustion camera at the beginning of

the disturbance. However, the temperature of the camera will be increased over the initial value as soon as the fuel begins to burn, which characterizes a system with inverse response where an increasing of the fuel feed causes, in the beginning, a decreasing of the temperature followed by an increasing of the temperature over the initial value. A physical explanation for this phenomenon is that the added mixture intensifies the burning duty.

The convolution of signals is clearly understood analyzing a system composed by two tanks, with continuous flow, where the input flow occurs at the top and the output flow from the bottom and discharge is proportional to the square root of the level of the respective tank. In this example, the discharge of the first tank is the input of the second one. In this system, the transfer function which relates the level of the second tank as a function of the feed flow of the first tank shows a second order behavior due to two sequential effects. The disturbance in the feed flow of the first tank implies in a dynamic change of the level of the first tank which is strongly related with its discharge (effect 1). Hence the modification in the discharge of the first tank will cause a dynamic change in the level of the second tank (effect 2).

Although these two different effects can occur in practice, it is possible to convert mathematically two sequential effects in a summation of two effects using a Heaviside expansion as follows:

$$\frac{K_1}{\tau_1 \cdot s + 1} \cdot \frac{K_2}{\tau_2 \cdot s + 1} = \frac{A}{\tau_1 \cdot s + 1} + \frac{B}{\tau_2 \cdot s + 1} \quad (4)$$

$$\frac{K_1}{\tau_1 \cdot s + 1} \cdot \frac{K_2}{\tau_2 \cdot s + 1} = \frac{A \cdot (\tau_2 \cdot s + 1)}{\tau_1 \cdot s + 1} + \frac{B \cdot (\tau_1 \cdot s + 1)}{\tau_2 \cdot s + 1} \quad (5)$$

$$\begin{bmatrix} K_1 \cdot K_2 \\ 0 \end{bmatrix} = \begin{bmatrix} A + B \\ A \cdot \tau_2 + B \cdot \tau_1 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \frac{K_1 \cdot K_2}{1 - \frac{\tau_2}{\tau_1}} \\ \frac{K_1 \cdot K_2}{1 - \frac{\tau_1}{\tau_2}} \end{bmatrix} \quad (7)$$

As shown in equation , it is possible to represent a convolution of two signals (1 and 2) as a summation of two signals of first order using a right set of apparent gains (A and B), since the ratio between time constants of these two original signals is not unitary. In cases where such relation is equal to one it is not possible to represent the convolution of signals through the summation of first order functions. In these cases, it is necessary that at least one of the base functions is of the second order. Although this condition is necessary, it will be assumed, for while, that all higher order functions can be represented as a summation of functions of first order and the case of multiple equal poles will be treated after other considerations to allow a better understanding.

### 2.3 Final formulation of the Model

Considering previous explanations the final formulation of the model can be represented as shown in equation assuming its vectorial expression and considering as many states  $x$  as necessary to represent all outputs and its respective orders. For instance, if the system has two outputs, being the first one of second order and the second one of third order, the system will be represented using five states, two of these states to represent the first output and the other three states to represent the second output. This assumption can take the proposed model to belong much more states than the original model based on first principles, since that each output conserves its own group of states and each state refer itself to a specific output. At first view, it seems to turn the problem more complex than the original, however these conversion show an important advantage to applications of real-time optimization. Due to the formulation of this model, it is possible to evaluate the solution of the model in a analytical way instead through a numerical integration. Equation 8 shows the analytical solution for the equations of the model shown in equation .

$$x_f = (x_i - ss(u_f)) \cdot e^{(p \cdot (t_f - t_i))} + ss(u_f) \quad (8)$$

Where index  $f$  represents the final value and index  $i$  the initial value.

### 3 Methodology for Model Conversion

The methodology shown until now cannot be considered as an identification methodology since there is no proposition of perturbation tests, statistical evaluation or treatment of process data. Every data used in this model conversion must be acquired from another model which describes the process with an acceptable quality. On the other hand, the method cannot be considered as an order reduction method since that, sometimes, the number of states of the resulting model is higher than the original model. The proposed method consists of a way to obtain a dynamic nonlinear process model quicker to be evaluated and useful in real-time optimization.

#### 3.1 Multidimensional Space

The conversion method needs the user to perform perturbations in the original model to generate trends of data, on which the conversion will be made. The perturbations must take the process model to specific regions in order to build a rectangular grid of steady-states. Such regions are determined based on the range of each input variable. In other words, a space with  $nu$  dimension, where  $nu$  is the number of inputs, must be created and sets of standard values must be established for each input. Using combinations of these sets of values to the different inputs, a rectangular grid of steady-states must be established. For instance, it can be considered a system with two inputs  $u_1$  and  $u_2$ , a set of standard values to be considered can be  $u_1 = [500, 750, 1000]$  and  $u_2 = [0.1, 0.5, 1]$ . In this case, there is three points established for each input and a rectangular grid of points that must be built is shown on table 1:

**Table 1:** Example of points of the rectangular grid

Points	$u_1$	$u_2$
1	500	0.1
2	750	0.1
3	1000	0.1
4	500	0.5
5	750	0.5
6	1000	0.5
7	500	1.0
8	750	1.0
9	1000	1.0

:

Of course, it is not always possible to obtain a rectangular grid composed by attainable steady-states. Sometimes, some points generated by the

combination of standard values of inputs are not feasible to be evaluated by the model and, until now, this is a real limitation of this method.

Since the points are established, it is necessary to make the original model generate results of transitions among such points. However it is very important to establish the best sequence of points to perform the transition. Considering that the objective is to generate multivariable dynamic model, it is very important to consider a most significant direction of perturbation. It means that it is necessary a previous knowledge about the behavior of the plant to guarantee that the perturbation direction will show all details of the model. In general, the perturbation direction chosen is the input direction of the transmission zero or that one which presents the biggest singular value. The direction determined in this way is called “best input direction”.

Of course, it is not very probable that the chosen perturbation direction can be observed using transitions among the points previously determined, however this direction corresponds to a best perturbation direction and other similar directions can bring a significant information even though they are not collinear to the best input direction. The similarity criterion to determine the best transition directions is the dot product between the best input direction determined and the possible direction determined by the transition among the points previously established. It is also necessary to normalize the directions before the comparison to guarantee that the dot product calculated depends only on the angle between the vectors. An algorithm determines the new sequence of the previous points determined performing successive changes in the order of the points, evaluating the dot product on each change searching for the best results. For instance, using the example of table 1 and considering the best input direction  $[0.7071, 0.7071]$ , the best sequence to the group of points generated is shown on table 2:

**Table 2:** Points of table 1 organized following direction  $[0.7071, 0.7071]$

Points	u1	u2
1	500	0.1
8	750	1.0
4	500	0.5
9	1000	1.0
2	750	0.1
6	1000	0.5

5	750	0.5
3	1000	0.1
7	500	1.0

Table 2 indicates the best sequence of perturbations (i.e. it is better to change from point 1 to point 8 than change from point 1 to point 2, following the criterion of best input direction [0.7071, 0.7071]).

### 3.2 Model Parameter Determination

Since the best sequence of perturbations has been determined, the simulation of the original model can be performed. It is necessary to establish a simulation time for each perturbation long enough to the system achieve the steady-state on each interval. Another condition to perform the model conversion is that, after the end of the sequence of perturbation has been reached, it is necessary to run back the simulation through the same perturbation points in an inverse order. The reason why this return is done will be explained in the section "Matching Related States".

Defining,  $Y$  as the trend of outputs generated, where each line contains simulated points in time and each column contains values for each output,  $U$  as the set of inputs produced to the simulation organized in the same way of the outputs,  $T$  vector of time, all points collected from simulation with a constant sample period  $t_s$  and on each  $n$  sample periods a new perturbation must be performed. In this way, the data trends generated must be divided in multiple interval of perturbation containing the following information:

- Starting output point
- Values of inputs (must be constant for)
- Response of the perturbation getting the steady-state

For the  $k$ -th interval the trend of data selected must be:

$$y_{Ak} = \begin{bmatrix} Y_{(k-1)n+1,1} & \cdots & Y_{(k-1)n+1,J} \\ \vdots & \ddots & \vdots \\ Y_{kn,1} & \cdots & Y_{kn,J} \end{bmatrix} \quad (8)$$

$$u_{Ak} = \begin{bmatrix} U_{(k-1)n+1,1} & \cdots & U_{(k-1)n+1,J} \\ \vdots & \ddots & \vdots \\ U_{kn,1} & \cdots & U_{kn,J} \end{bmatrix} \quad (9)$$

$$t_{Ak} = \begin{bmatrix} T_{(k-1)n+1} \\ \vdots \\ T_{kn} \end{bmatrix} \quad (10)$$

Since the previous conditions has been satisfied, it is possible to proceed with model conversion. The first step to identify model parameter needs the removal of initial bias from each output data set in order to make the initial point of each trend to be equal to zero, as follows:

$$y_k = y_{Ak} - y_{Ak_1} \quad (11)$$

In the same way, time to be considered at this point needs to be initialized in zero, as follows:

$$t_k = t_{Ak} - t_{Ak_1} \quad (12)$$

The idea of this methodology is to determine multiple models spread on the multidimensional space of input variables, and interpolate these models to build a nonlinear model that represents the process. The local models are obtained from each transition performed in the original model. The requirement for the model identified using the selected data set is that, for any kind of response, the identified model have to be a summation of first order models. In addition models identified must allow a specification of an order required to guarantee that all models obtained will have the same number of states for each output. The method developed to perform such identification consists of a progressive algorithm which successively increases the order of the system.

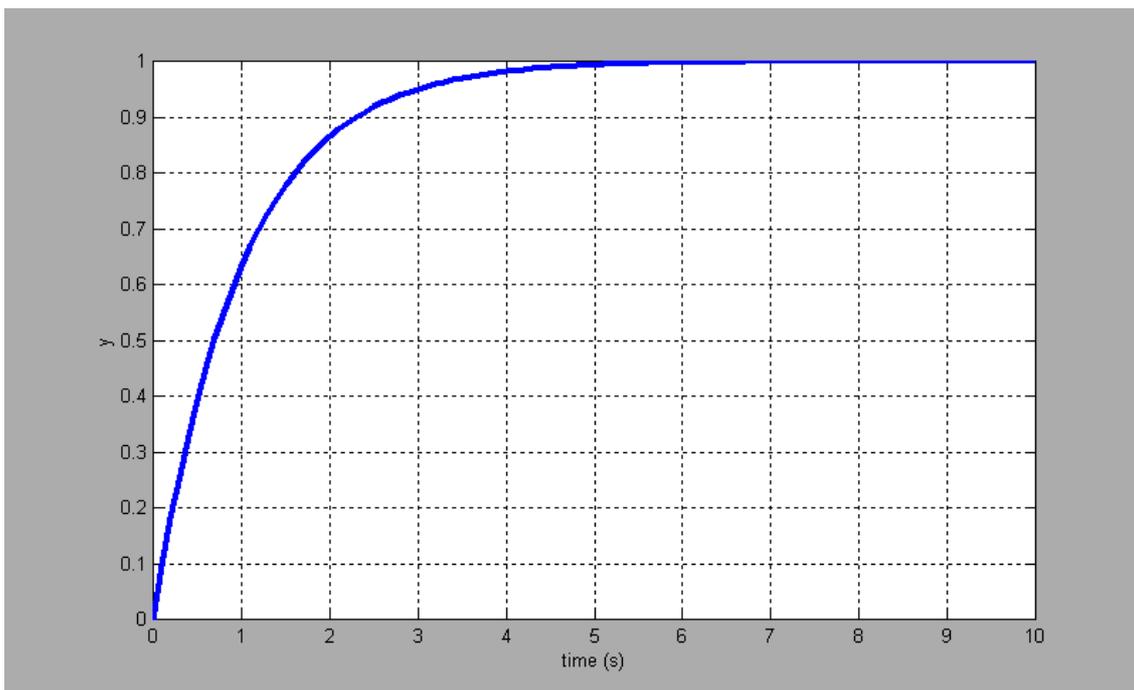
### 3.3 **Local Model Determination**

A previous explanation is necessary to understand how the algorithm starts and why its progression is made in the proposed way. In this subsection, it will be detailed how to perform the determination of a local model based on the data of the original model.

Consider a response, shown in figure 3, produced for a unitary step produced in a system represented by the following transfer function:

$$G_1(s) = \frac{K}{\tau s + 1} \quad (13)$$

Where  $K=1$  and  $\tau = 1$ .



**Figure 3:** Response of a first order function for  $K = 1$  and  $\tau = 1$

The analytic solution for the step response of is represented by:

$$y_1 = K \left( 1 - e^{\left(-\frac{t}{\tau}\right)} \right) \quad (14)$$

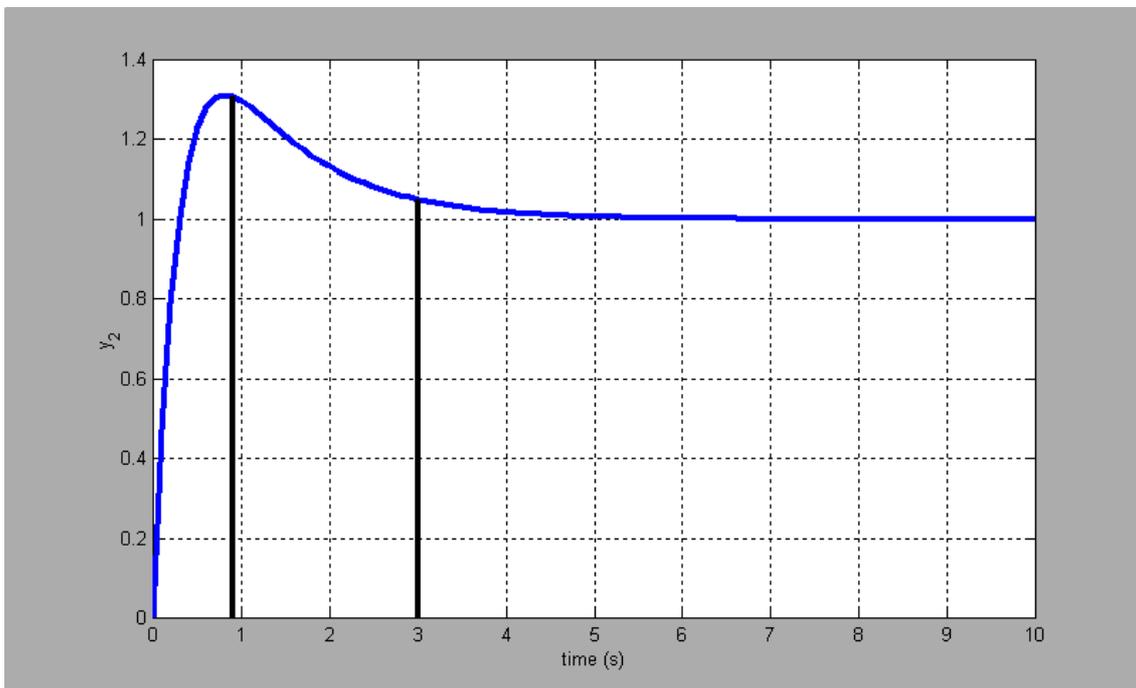
It is observed that the changes in the output of the process occurs during a specific time (settling time) and, afterwards, there are no significant changes in the value of the output. This behavior happened because the time response can be represented by an exponential approximation of the function to a constant value ( $K$ ). In other words it is possible to say that a first order function produces dynamic effects on the response while the exponential term of its representation have a significant value and, afterwards, a constant value is

assumed. Here in this work it will be assumed that when  $t > 3\tau$  the dynamic effects of the function will be ended.

Consider now a response, shown in figure 4, produced for a unitary step produced in a system represented by the following transfer function:

$$G_2(s) = \frac{K_1}{\tau_1 s + 1} + \frac{K_2}{\tau_2 s + 1} \quad (15)$$

With  $K_1 = -1$ ,  $\tau_1 = 1$ ,  $K_2 = 2$  and  $\tau_2 = 0.3$ .



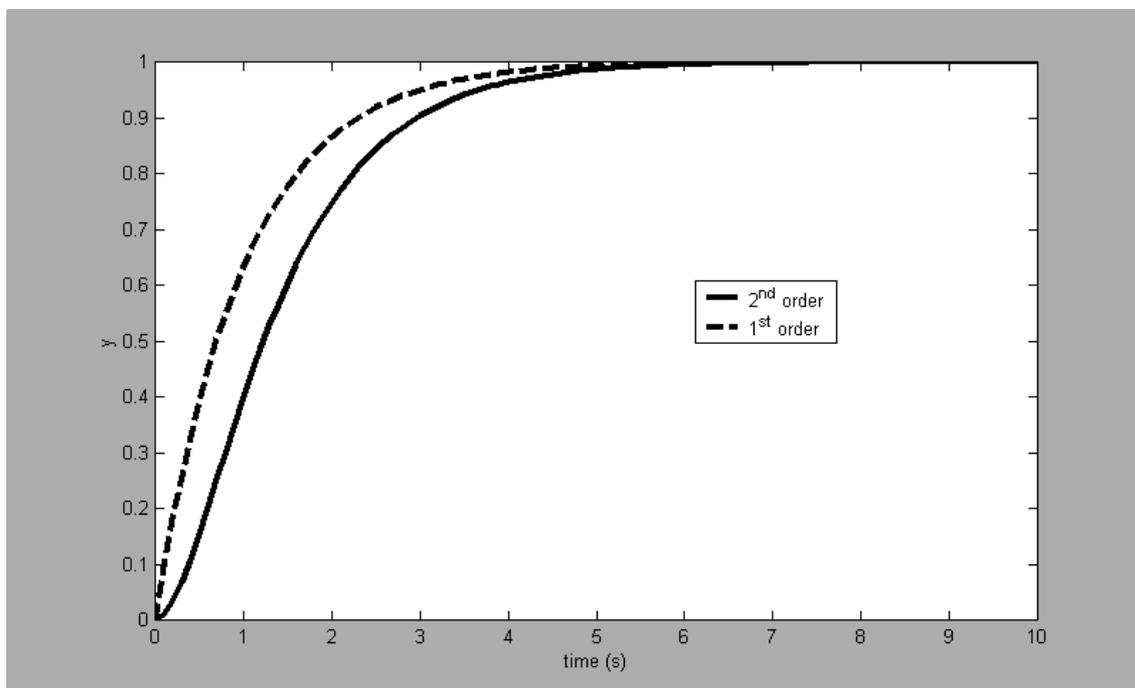
**Figure 4:** Summation of response of a first order function ( $K_1 = -1$ ,  $\tau_1 = 1$ ,  $K_2 = 2$  and  $\tau_2 = 0.3$ )

The analytic solution for the step response of is represented by:

$$y_2 = K_1 \left(1 - e^{\left(-\frac{t}{\tau_1}\right)}\right) + K_2 \left(1 - e^{\left(-\frac{t}{\tau_2}\right)}\right) \quad (16)$$

Figure 4 shows the response of the analytic system represented on equation . The vertical lines were placed at times 0.9 s and 3 s which are respectively  $3\tau_2$  and  $3\tau_1$ . Note that the response is commanded basically for the gain  $K_2$  associated to  $\tau_2$  in the beginning of the response and, after  $t > 3\tau_2$  the response starts to be considered by the gain  $K_1$  associated to  $\tau_1$  until a time  $t = 3\tau_1$ , when the response starts to be constant ( $K_1 + K_2$ ). It is easy to see that the behavior imposed by a gain associated to a dynamic base is important while

the response time is smaller than the time related to the respective dynamic base ( $t = 3\tau$ ). Therefore, if the variations of the behavior of the response of the system are observed, it is possible to determine the important dynamic bases which describe the dynamic response. However, it is not enough to observe the changes in the derivative signal of the response, since they will not always exist, as shown in figure 5 where the response of a second order system without inverse response or overshoot is illustrated. In such system, it is possible to observe that although there is not an inversion in the signal of the derivative of the response, the second order response shows a significant distance from the first order response with the same settling time, which indicates the presence of another dynamic base.



**Figure 5:** Summation of response of a first order function ( $K_1=-1$ ,  $\tau_1 = 1$ ,  $K_2=2$  and  $\tau_2 = 0.3$ )

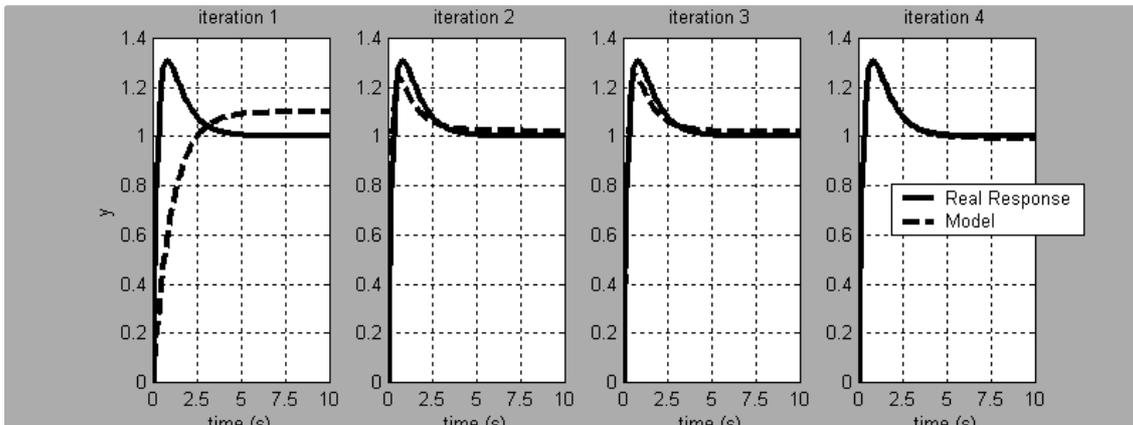
Therefore, an iterative method for model determination based on a response of a system was determined. The method starts performing a poor approximation of the response using a first order model with the same settling time of the shown response. The determination of such system is made using the observed gain of the original system and using a time constant equal to the settling time of the original response, considering such settling time as the time when the system achieves a band around the steady-state smaller than 5% of the absolute value of the gain. Once that such first order system has been obtained, it is possible to determine other dynamic bases to improve the representation, adding other first order transfer functions which approximate in

a better way the model. The determination of the following time constant is based in the biggest difference between the original response and the model obtained. The intention with this operation is to describe the response as best as possible, eliminating the biggest differences. Since the new dynamic base has been determined, it is possible to find the associated gains solving the linear system given by:

$$K_{1est} \left( 1 - e^{\left( -\frac{t_{pt1}}{\tau_1} \right)} \right) + K_2 \left( 1 - e^{\left( -\frac{t_{pt1}}{\tau_2} \right)} \right) = y_{pt1} \quad (17)$$

$$K_{1est} \left( 1 - e^{\left( -\frac{t_{pt2}}{\tau_1} \right)} \right) + K_2 \left( 1 - e^{\left( -\frac{t_{pt2}}{\tau_2} \right)} \right) = y_{pt2} \quad (18)$$

If the difference between response and model persists, a new dynamic base can be determined based on the biggest remaining difference, and solving a new linear system again. Figure 6 shows four iterations to determine a representative model of a signal produced by a step in a second order transfer function of equation . Table 3 shows the gains K and time constants  $\tau$  obtained at each iteration step of the determination algorithm.



**Figure 6:** Model obtained in each one of the four iterations performed

**Table 3:** Gains and Time Constants obtained during the iterative determination of the model shown in figure 6

Parameter		Iterations			
		1	2	3	4
K	1	1.0997	1.4765	-0.7798	-0.1718
	2		-0.4541	2.3266	0.7479
	3			-0.5281	1.4695
	4				-1.0526
$\tau$ (s)	1	1.0333	0.1667	0.0667	0.0667
	2		1.0333	0.1667	0.1667
	3			1.0333	0.3667

	4				1.0333
--	---	--	--	--	--------

The approximation performed to describe the response shows that it was necessary a fourth order system to represent a response produced by a second order which seems to be not efficient. However, the second part of the algorithm carries out the order reduction of the model. It is only necessary to keep in the model significant portions of the model, being the others discarded. The selection of a discarded portion is done considering the gain obtained in the last approximation. In general, very little gains turn a portion not significant and, therefore, it can discard the respective portion. Figure 7 shows the responses obtained during the order reduction of the previous model obtained. The Gains and time constants can be seen in table 4, which shows the gains and time constants obtained during order reduction.

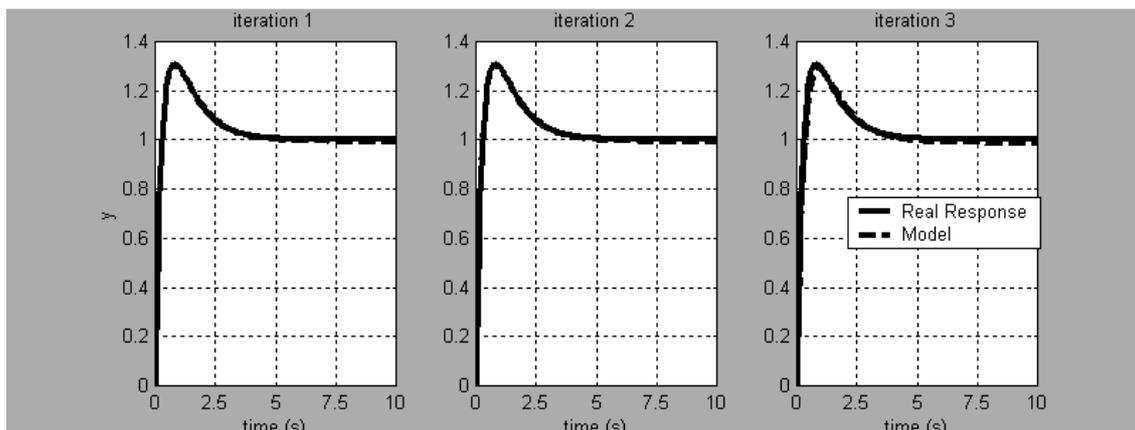


Figure 7: Model obtained in each one of the four iterations performed

Table 4: Gains and Time Constants obtained during the order reduction of the model shown in figure 7

Parameter		Iterations		
		1	2	3
K	1	-0.1718	0.5189	2.1512
	2	0.7479	1.5362	-1.1637
	3	1.4695	-1.0628	
	4	-1.0526		
$\tau$ (s)	1	0.0667	0.1667	0.3667
	2	0.1667	0.3667	1.0333
	3	0.3667	1.0333	
	4	1.0333		

The algorithm shown here demonstrates a new way to “adjust” a model to match with a time response generated by another model. It is important to point out that such algorithm is not an identification method since it do not works with noisy signals. It is only a way to convert a step response model into a transfer function model or state space model composed by summation of first order models. The advantage of using this method is the absence of an optimization problem which implies in a quick way to convert a model. It is important to note also that, even though the reduction step is not performed, the approximation of the original signal can be as good as the order of the model determined. In other words, it is possible to control the approximation order as much as desired to obtain several models of multiple responses always with the same order.

### 3.4 Matching Related States

The determination of the global model depends on a good identification of the local models as previously explained. As explained in previous sections, the sequence of perturbation is simulated in a specified order following a required preferential direction. After such perturbations has been simulated the inverse way is followed, and all points of the orthogonal grid are simulated again beginning in the last point of the set and finishing at the first one. The reason for developing this second sequence of perturbation is to determine the best interpolation order for the states of the model, where the first simulation is used to determine the local model and the second one to determine the best interpolation of this model with the next one.

Demonstrating the interpolation method, consider the situation to interpolate the models obtained to the interval  $k$  and  $k+1$ , for models with order  $n$ . In this situation, there is one model for each interval given by:

$$\bar{y}_{k-1 \rightarrow k} = \sum_{i=1}^n K_{k,i} \left( 1 - e^{\left( -\frac{t_k}{\tau_{k,i}} \right)} \right) \quad (19)$$

$$\bar{y}_{k \rightarrow k+1} = \sum_{i=1}^n K_{k+1,i} \left( 1 - e^{\left( -\frac{t_{k+1}}{\tau_{k+1,i}} \right)} \right) \quad (20)$$

The problem here is to determine what is the best relation among the gains  $K_{k,i}$  and  $K_{k+1,i}$  since the final point of the interval  $k$  (point  $k$ ) is the initial point of the interval  $k+1$  and the transition from the final point of interval  $k+1$  to

the initial point of the same interval (inverse simulation) must follow the  $\tau_{k|i}$  instead of  $\tau_{k+1|i}$ . This conclusion comes from the precept that the dynamic of a transition depends only on the end point of the transition. Hence the model for the inverse transition is given by

$$\bar{y}_{k+1 \rightarrow k} = \sum_{i=1}^n K_{k+1,j} \left( 1 - e^{-\frac{t_{k+1}}{\tau_{k,i}}} \right) \quad \text{for } j = f(i) \quad (21)$$

Note that, in this situation, the gain associated to each  $\tau_{k|i}$  is not the respective  $K_{k+1|i}$ , but  $K_{k+1|j}$  where  $j$  is function of  $i$ . This happens because, not necessarily, the  $i$ -th gain obtained in the local model determination  $k+1$  will represent the best approximation of transition  $\bar{y}_{k+1 \rightarrow k}$  if associated with  $\tau_{k|i}$ . Remember that the local model determination is an approximation for the full model and the parameters  $K$  and  $\tau$  are obtained in order to describe the time response and, not necessarily, are related with real summation of physical effects. Hence, it is necessary to find the best relation among gains obtained in the interval  $k$  with those ones obtained in the interval  $k+1$ . The answer for this problem is in the discrete optimization of the error between the trajectory of the response of the inverse simulation and the prediction given by the model as follows:

$$\min_{\text{order of } K_{k+1|j}} S = \sum (y_{k+1 \rightarrow k} - \bar{y}_{k+1 \rightarrow k})^2 \quad (22)$$

Since the optimization takes to the best order of  $K_{k+1|j}$  it is possible to associate each  $K_{k+1|j}$  to each  $\tau_{k|i}$  and, hence, to each  $K_{k|i}$ . Based on this association, it is possible to perform the next step in the model construction, i.e., the determination of steady-state functions.

### 3.5 Determination of Steady-State Functions

Considering the states of the model are correctly matched, it is possible to build the steady-state functions "ss" for each state of equation . As demonstrated before, each transfer function of first order generates a state of the model. The first step is to determine the value of steady states functions at the specific steady-states reached by the original model. The procedure to find such values consists of attribute an initial value of steady-states of the initial

starting steady-state and progress determining the steady-states values for the subsequent steady-state reached during the simulation. Since the gains for each interval is known and the association of gains of two sequential models specified, it is possible to determine the steady states value adding the related gains of sequential models, as follows:

$$SS_{k+1i} = SS_{kli} + K_{k+1j(i)} \quad \text{for } i = 1, \dots, n \quad (23)$$

By definition:

$$SS_{0!1} = SS_0 \quad (24)$$

$$SS_{0!i} = 0 \quad \text{for } i = 2, \dots, n \quad (25)$$

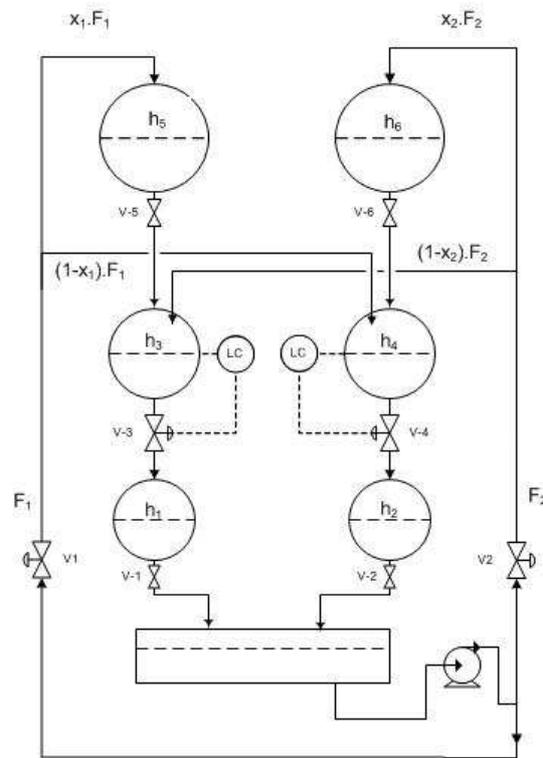
The determination of the poles  $p$  of the system can be made straight using the values of time constants obtained, as follows:

$$p_{kli} = -\frac{1}{\tau_{k|i}} \quad \text{for } i = 1, \dots, n \quad (26)$$

Since the steady-states and poles are well specified on the points of the rectangular grid, it is necessary to interpolate them to specify a function capable to describe all region of operation of the model. Although techniques of multidimensional splines are well developed, in this work the linear interpolation has been used, due to its simplicity and guarantee that all interpolated values are inside of the extreme values. Note that in cases where the poles is near to zero high order interpolation can produces positive poles , which is undesirable.

## 4 | Case Study

The example chosen to apply this method was the six spherical tanks plant which is shown in figure **8**



**Figure 8:** Schematical representation of the sextuple spherical tank plant

This plant is composed by a set of six spherical level tanks connected among them in a special sequence to promote different behaviors. The six tanks numbered from 1 to 6 are displayed as shown on figure. The tanks are placed in two columns with three tanks each one. The tanks 1, 3 and 5 are placed on the left column and the tanks 2, 4 and 6 on the right column. The system feed is composed by two streams  $F_1$  and  $F_2$  which takes the water from the bottom buffer tank and distribute into the four tanks. The feed flow  $F_1$  is divided in two streams through a fraction  $X_1$  where the flow  $F_1 \cdot X_1$  is fed over the tank 5 while the rest of this flow is sent to the tank 4. In other way, the feed flow  $F_2$  is divided in two streams through the fraction  $X_2$  where the flow  $F_2 \cdot X_2$  is sent to the tank 6 while the rest of this stream is loaded in the tank 3. In each column of tanks the bottom flows are sent to the subsequent tank immediately under that ones, or in other words, the bottom flow of tank 5 is fed on tank 3 and the bottom flow of tank 3 is sent to tank 1. In the same way, the bottom flow of tank 6 is fed on tank 4 and the bottom flow of tank 4 is sent to tank 2. The bottom flows of the system are sent to the reservoir to be used again through the flows  $F_1$  and  $F_2$ .

The first principle model of this plant can be represented by:

$$\frac{dh_1}{dt} = \frac{C_3 \cdot \sqrt{h_3} - C_1 \cdot \sqrt{h_1}}{\pi \cdot h_1 \cdot (D_1 - h_1)} \quad (27)$$

$$\frac{dh_2}{dt} = \frac{C_4 \cdot \sqrt{h_4} - C_2 \cdot \sqrt{h_2}}{\pi \cdot h_2 \cdot (D_2 - h_2)} \quad (28)$$

$$\frac{dh_3}{dt} = \frac{F_2 \cdot (1 - X_2) + C_5 \cdot \sqrt{h_5} - C_3 \cdot \sqrt{h_3}}{\pi \cdot h_3 \cdot (D_3 - h_3)} \quad (29)$$

$$\frac{dh_4}{dt} = \frac{F_1 \cdot (1 - X_1) + C_6 \cdot \sqrt{h_6} - C_4 \cdot \sqrt{h_4}}{\pi \cdot h_4 \cdot (D_4 - h_4)} \quad (30)$$

$$\frac{dh_5}{dt} = \frac{F_1 \cdot X_1 - C_5 \cdot \sqrt{h_5}}{\pi \cdot h_5 \cdot (D_5 - h_5)} \quad (31)$$

$$\frac{dh_6}{dt} = \frac{F_2 \cdot X_2 - C_6 \cdot \sqrt{h_6}}{\pi \cdot h_6 \cdot (D_6 - h_6)} \quad (32)$$

where  $C_i$  is the head loss of the discharge of  $i$ -th tank;  $D_i$  is the diameter of the  $i$ -th tank;  $h_i$  is the level of water in each tank;  $F_j$  is the feed flow and  $X_j$  is division fraction of the  $j$ -th flow.

The values used for constants of the model are given in table 5.

**Table 5:** Values of parameters of the model

Tank	Parameter	
	C (cm <sup>2.5</sup> /s)	D (cm)
1	16.0428	26
2	16.0428	26
3	16.0428	26
4	16.0428	26
5	16.0428	26
6	16.0428	26

There are two interesting characteristic of this plant. The first characteristic is the appearance of positive transmission zeros which occurs in a special situation when there is an increasing in one flow and a decreasing in another one. In this case, each one of this perturbations are felt by the process in different times and the occurrence of a perturbation with different signs in each element makes opposite effects in the response which can generate inverse response or overshooting. The second important characteristic of this

system is the presence of the inversion of the dominance of the manipulated variable on the level of the tanks. The presence of the division of the flows through the fractions  $X_1$  and  $X_2$  makes that each flow prevail on a specific column of tanks depending of the summation  $X_1+X_2$ . If this summation is bigger than the unit the left column is dominated by flow  $F_2$  and the right column by the flow  $F_1$ . Is the result of the summation is smaller than the unit the dominance of the system is inverse.

Such characteristics promote the appearance of too many different behaviors for each operation point and, therefore, the nonlinearities are evidenced.

## 5 Procedure for Model Conversion

Showing the application of this algorithm, the level of the tank 2 was chosen as output variable to perform the model conversion. The points used to build the rectangular grid were chosen in order to describe entire useful space of all variables and, therefore, three points of each feed flow and four points of each flow fraction were evaluated – a minimum point, a maximum point and one or two intermediate points. The steady-state input points chosen to compose the rectangular grid are show in table 6.

**Table 6:** Values of manipulated variables used to build the rectangular grid

Variable	Values			
	Minimum	Intermediate 1	Intermediate 2	Maximum
$F_1$ (cm <sup>3</sup> /x)	32	54	-	77
$F_2$	32	54	-	77
$X_1$	0.1	0.3	0.6	0.9
$X_2$	0.1	0.3	0.6	0.9

The set of values of input variables shown before originates rectangular grid containing 144 steady-state points, which are organized following the perturbation direction [-100 100 1 1]. Such direction was chosen considering the transmission zero of the system. It is important to point out that such direction is not a calculated zero direction, but an estimated zero direction. The reason why such direction was not precisely calculated is to attribute to the algorithm robustness in determining the model conversion, since that it is not necessary to provide a precise zero direction, but only a approximation based on the knowledge of the plant. Moreover, the direction chosen cannot be exactly

reproduced since the steady-state points are previously established and the perturbation direction is given by them.

The estimation of the search direction is made considering the response of the perturbation of different input variables on the level of tank 2 (the considered output variable). The generation of an inverse response or a overshooting in this system occurs even if the perturbation in the straight feed flow of tank 4 ( $F1*(1-X1)$ ) were opposed to the feed flow of tank 6 ( $F2*X2$ ), (i.e.  $F1*(1-X1)$  decreasing and  $F2*X2$  increasing). This behavior appears because the interest level ( $h_2$ ) is quickly affected by the first flow (dynamic of tank 2 and 4) while the effect of perturbation in  $F2*X2$  comes later (dynamic of tanks 2, 4 and 6). The intensity of each variation will give the characteristic of the response. If the module of the perturbation in  $F2*X2$  is bigger than the module in perturbation of  $F1*(1-X1)$ , than an inverse response occurs. If such module is smaller than the other an overshoot takes place. Following such considerations and the range of each variable the direction  $[F1 F2 X1 X2]=[-100 100 1 1]$  was chosen. The set of inputs and its responses on  $h_2$  is shown in figure 9.

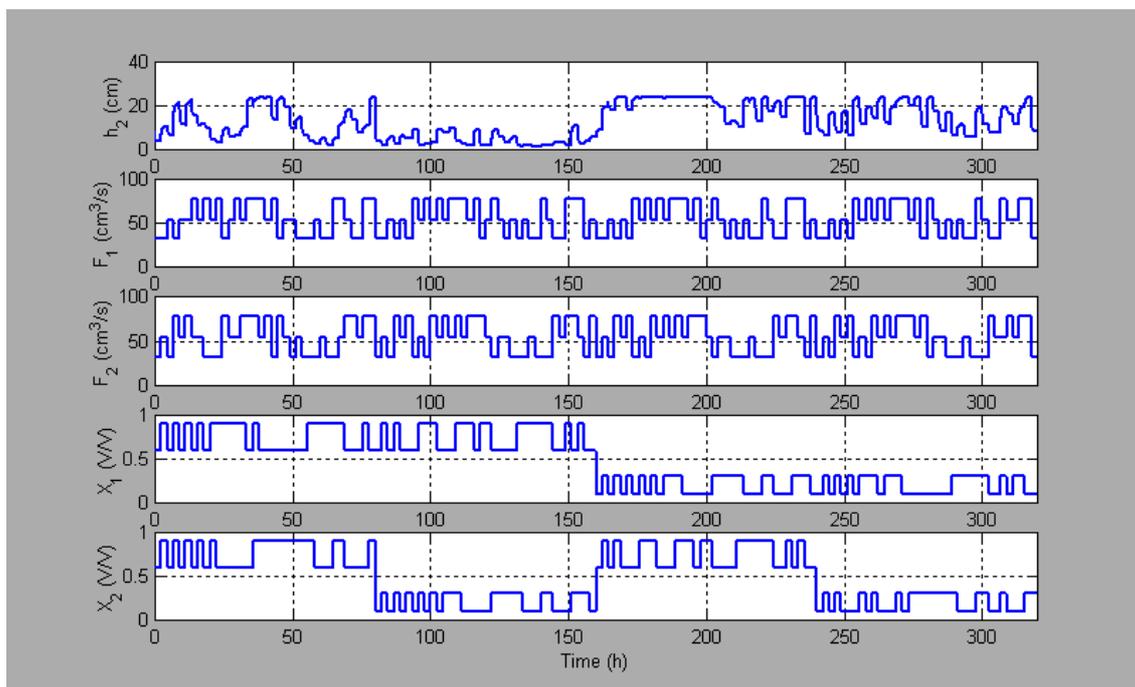


Figure 9: Identification data

As shown in figure 9 the first perturbation was forced to be that ones with higher values of  $X1$  and  $X2$ . The reason why this choice was done is the fact of inverse response and overshoot only occurs whether the levels of tank 6 is big because with low levels of this tank there is not a difference between the

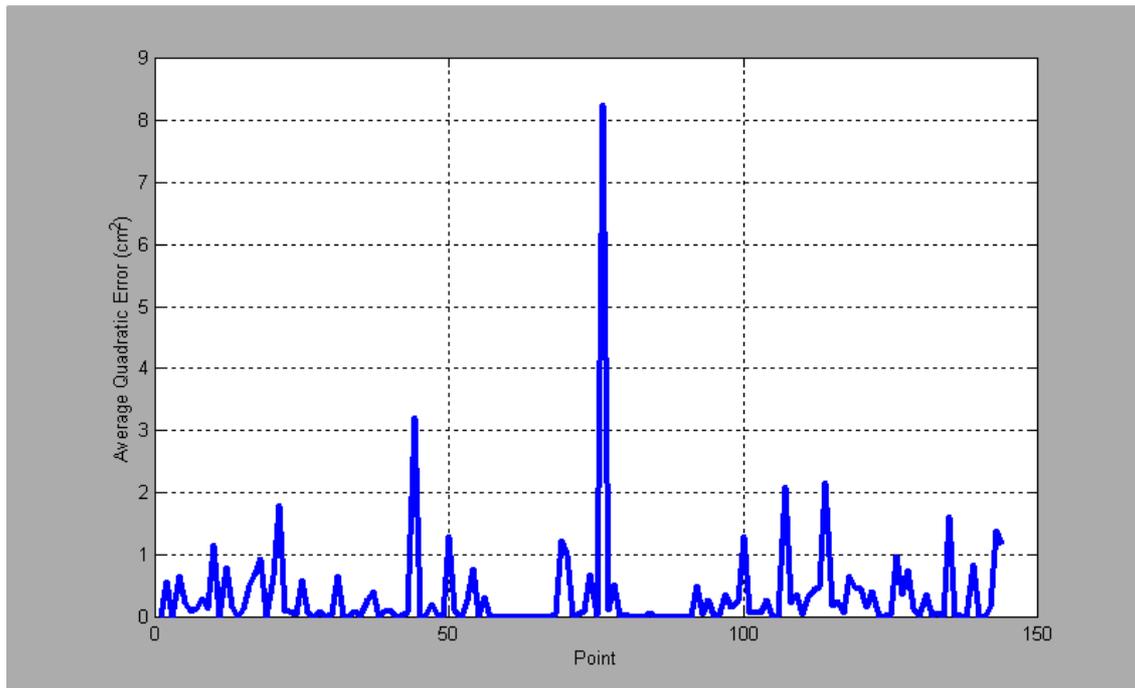
dynamics of the feeds and the response seems to be a simple response of first order. Hence the high level perturbations were separated from the others.

## 6 Results and Discussion

In order to evaluate the quality of obtained models, the criterion of Euclidian norm of the error in the data of the back perturbations was established as follows:

$$AQE = \frac{(y_{k+1 \rightarrow k} - \bar{y}_{k+1 \rightarrow k})^T \cdot (y_{k+1 \rightarrow k} - \bar{y}_{k+1 \rightarrow k})}{n} \quad (33)$$

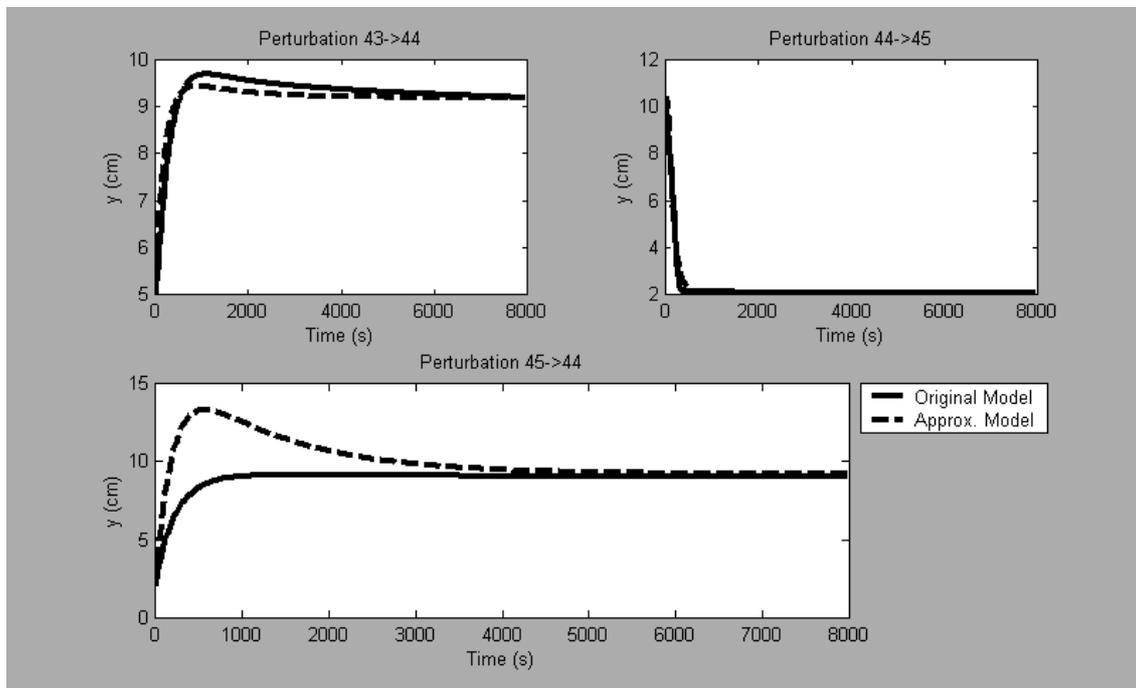
where AQE is Average Quadratic Error between the response of the original model and the approximate model. The reason why the Euclidian norm was not normalized is that every model approximated describes the same process in different operational regions and the absolute error is a most appropriate way to evaluate the quality of different models. Figure **10** shows the AQE for models identified for each perturbation performed.



**Figure 10:** Quality of the identified models for each perturbation

As shown in figure **10** two models present the most significant errors – models 44 and 76. The responses of such models are shown in figures **11** and which show the transition from the previous point to the point mentioned, the transition from the mentioned point to the next one and the back transition from

the last point to the mentioned point. In such figures, it can be clearly seen that the initial supposition which considered the dynamic of the transition dependent only on the final point is not observed (i.e. transition from point 43 to 44 presents a different dynamic of the transition from the point 45 to 44). This divergence between such dynamics is the main responsible for the bad quality of the models obtained in this two points shown. However this group of bad converted models represents a tiny set of models. Indeed, the biggest amount of converted models do not present such divergence and the final dynamic is really predominant as shown in figure **13**, or, whether such behavior is found its effect is not significant against the global model which can also be proven by the AQE shown in figure **10**.



**Figure 11:** Responses of the model around point 44

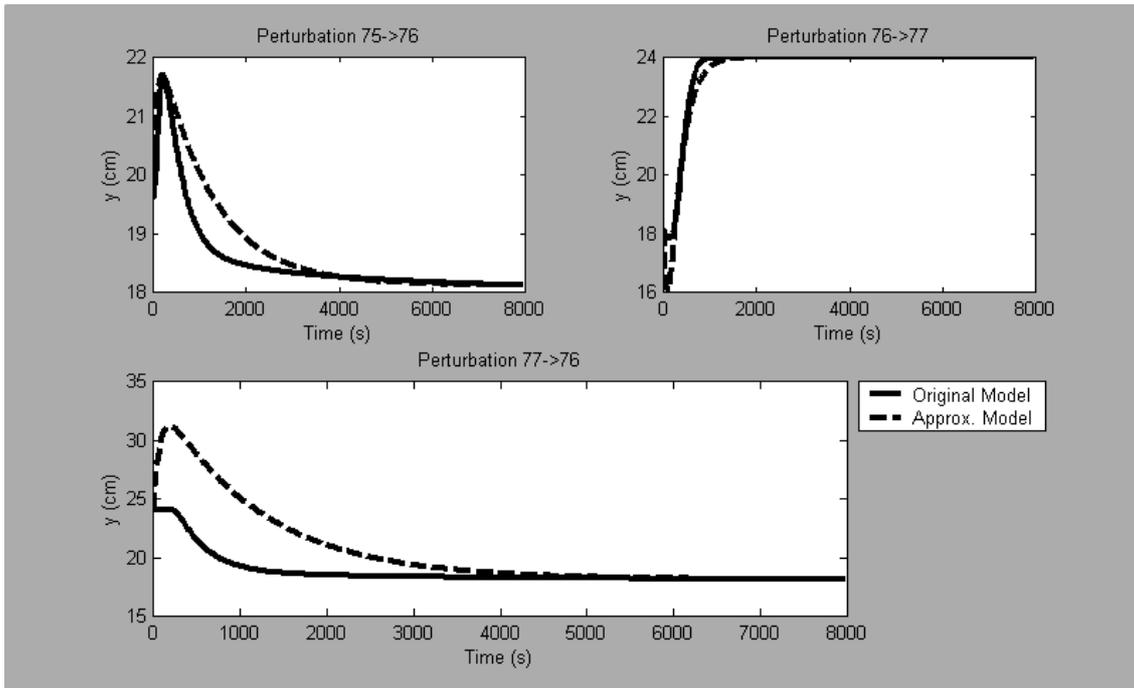


Figure 12: Responses of the model around point 76

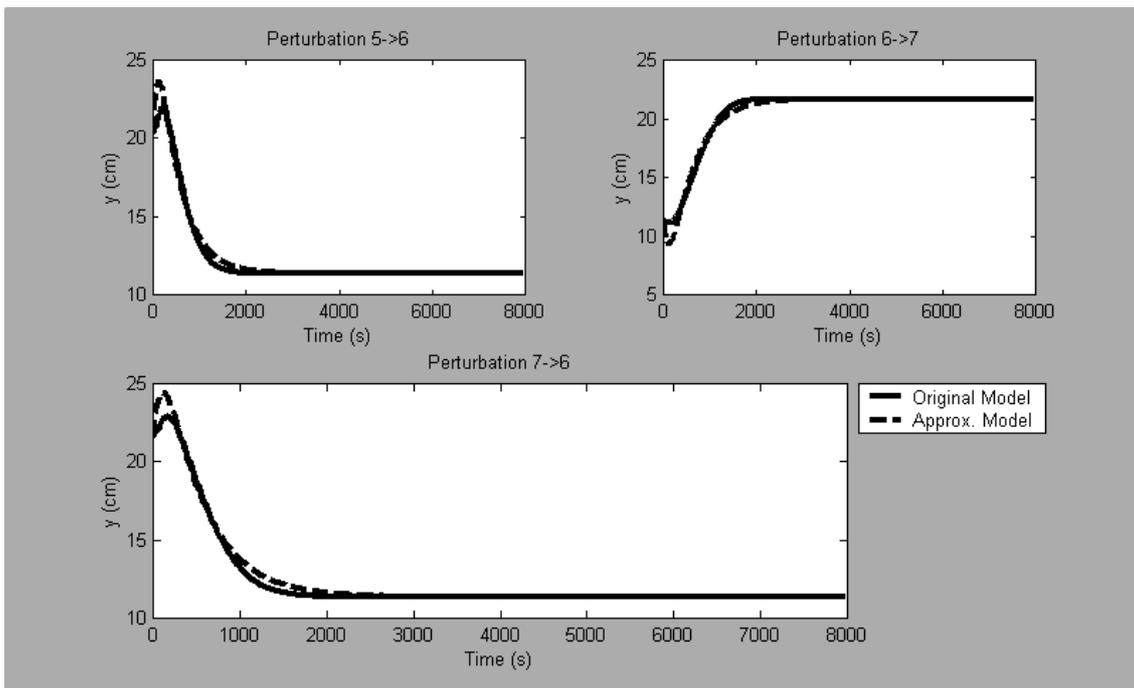
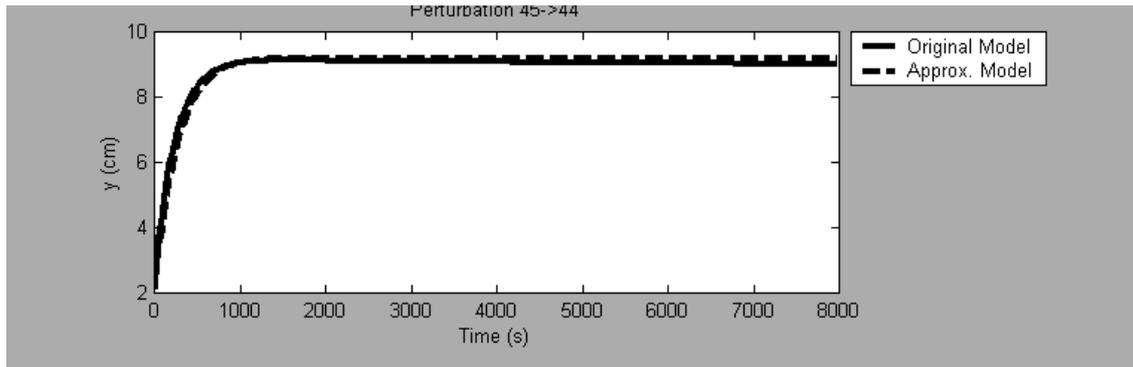


Figure 13: Responses of the model around point 76



**Figure 14:** Transition from point 45 to 44 using more adequate time constants

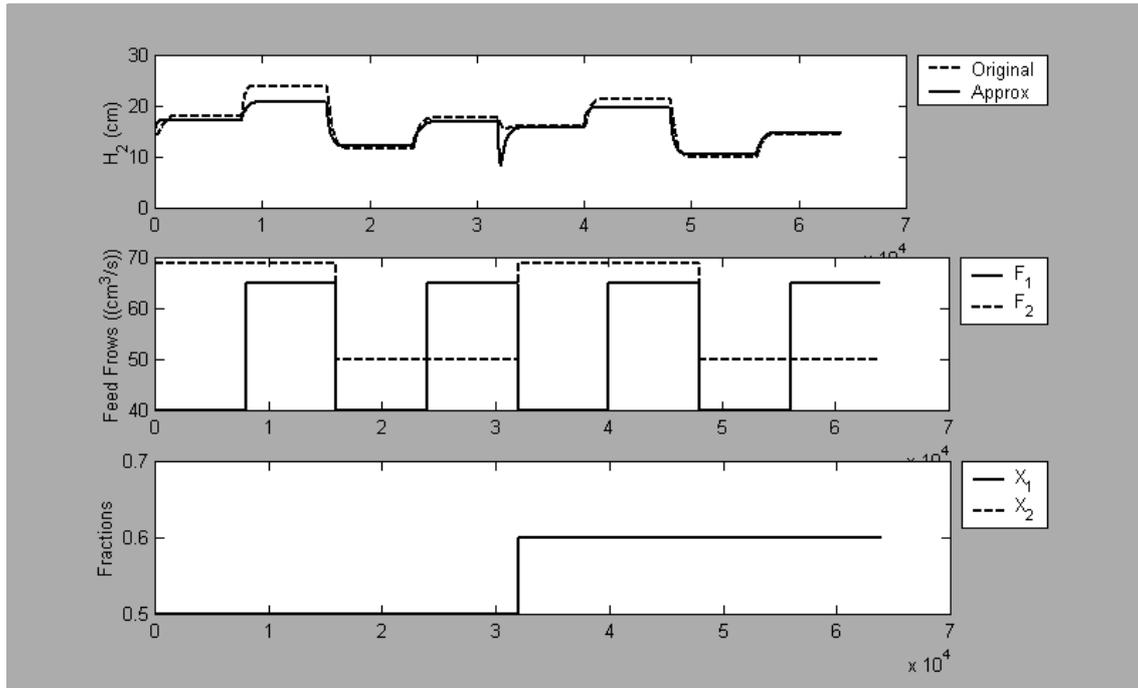
The situation shown in figure **11** can be improved by using the same set of gains identified in the transition from point 44 to 45 ( $K_1=-15.1597$  and  $K_2=8.0492$ ) but changing the original time constants identified for the point 44 ( $\tau_1 = 200.0$  and  $\tau_2 = 1193.3$ ) were changed for the time constants  $\tau_1 = 310.0$  and  $\tau_2 = 350$ . In this case the response obtained for transition from point 45 to 44, shown in figure **14**, indicates that an adjustment of the time constant is necessary depending on the direction of the transition.

The validation of the obtained model was done performing a sequence of perturbations in the full model composed by the set of linear models determined before. Table **7** presents the sequence of perturbations used in the validation.

**Table 7:** Values of manipulated variables used to build the rectangular grid

Variable	Perturbations used in validation							
	1	2	3	4	5	6	7	8
$F_1$ (cm <sup>3</sup> /x)	40	65	40	65	40	65	40	65
$F_2$ (cm <sup>3</sup> /x)	69	69	50	50	69	69	50	50
$X_1$	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6
$X_2$	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7

Figure **15** shows the response of the perturbations performed in the original model compared with the same perturbations performed in the converted model.



**Figure 15:** Perturbations performed to validate the model

The second and sixth perturbation performed show that the interpolation of steady-states in these two cases does not describe very well the steady-states achieved by the model. However, the other steady-states achieved were well described by the proposed model with a good quality. The dynamic of the fifth perturbation shows an overshoot bigger than the one produced original model. The other perturbations shows a good description of the original model. The time expense to integrate such perturbations for different solvers is shown in table 8. Note that the proposed algorithm is 45 times quicker than the quickest numerical solver. Considering that the model of the case studied is very small, the gain in time evaluation can be much bigger when applying such conversion for models of distillation columns, since the number of states of the conversion is proportional to the number of output variables.

**Table 8:** Values of manipulated variables used to build the rectangular grid

Solver	Time Spent (s)	Time Ratio
Approx	0.146	1.00
Ode23tb	7.096	48.5
Ode23	12.606	86.1
Ode113	19.705	134.6
Ode15s	6.642	45.4
Ode23s	11.406	77.9
Ode23t	7.441	50.1
Ode45	10.582	126.9

## 7 Conclusions

The results shown in this chapter presents an alternative model to be applied to Dynamic Real-time Optimization. Although it seems to be a quicker model to be evaluated, it still needs certain evolutions to be applied. In this final consideration it will be shown the advantages and disadvantages of using such kind of model for process optimization, pointing out the reasons why its application is promising although it is still not possible in some cases.

The results show that the initial supposition which establishes that the dynamic of a transition depends only on the final point is not completely right. Indeed, the dynamic of the transition depends on the lines of attraction shown in figure **2**, however the consideration of constant slope for each line for all space is not a correct consideration since the dynamic depends also on the initial point, from where, the transition starts as shown in figures **11** and . It is easy to prove that it is possible to find more appropriate time constants by modifying them manually as performed in figure **14**, which reinforce the theoretical that the dynamics depends on the dynamic of the ending point and the direction of the perturbation. Hence, the right formulation of equation must be proposed as show in equation :

$$\frac{dx}{dt} = p(x_0, u) \cdot (x - ss(u)) \quad (34)$$

However, this thesis could not present a way to determine such dependence of the dynamic with the direction of the transition, since the ordination of the gains and the association of them point after point uses the value of the time constants of each point to be performed. Therefore, if a change of the time constants is necessary following a given direction the method to determine the relation among states have to be reformulated

The same mismatch of dynamic behavior is found in the simulation to validate the model in the fifth perturbation where the dynamic presented by the approximated model is much more accentuated. Fortunately, the dynamic responses presented by the approximated model shows always a much more expressive inverse response or an overshoot the original which turns this model applicable for control and real time optimization, but with a probable decreasing in the quality of the results. However it is not possible to guarantee that this kind

of behavior will be always verified, and it is possible that in some cases the intensity of inverse responses and overshoots are smaller than the original, which can bring real problems for control and optimization.

The steady-states of the model could be well represented by the linear interpolation of the steady-states determined for the approximated model, except in little set of regions. It was not necessary to increase the order of the interpolation size; such operation could increase the computational time expenses without bringing any benefit for the final result.

As shown, this model approach is still not an applicable way to convert a nonlinear model into a simplified shape, however few developments in its shape can bring a new way to model industrial processes. Such models can be helpful for application in real time optimization considering the gain obtained in the computational time evaluating such model in comparison with first principle models.

The methods presented here show also others limitations since integrating and unstable models cannot be represented. In such cases the approach supposes a gain and a dynamic that can represent the response to perturbation performed but it is a unreal representation. Other important limitation is the discretization of the time constants. Considering that the time constants are given as a fraction of the settling time and such parameter can only be determined in a discrete way (i.e. settling time is determine on a point of the time response) the  $\tau$  constant will be dependent of discretization. Therefore it is necessary to determine a sample time capable to describe the response with good approximation but not too small such that increases the computational efforts. Heuristically, it was used 400 point for each perturbation which was appropriate in this case.

A quite interesting result of this work, which was not its objective but arises as a secondary method, is the obtaining of a new method for conversion of step response models into state-space models. Such method does not uses any kind of optimization in the conversion and can be totally adjusted and controlled by the user. However the method is restricted to obtain models with real poles. The researches for obtaining complex poles using the same philosophy but there was not a significant result to be shown.

## 8 | References

Fernandes, P. R. B. (2006). The Input-Parameterized Linearization around the Equilibrium Manifold Approach to Modeling and Identification. Dortmund, Universidade de Dortmund. **Doutorado:** 180.

Hedengren, J. D. and T. F. Edgar (2005). Order Reduction Of Large Scale DAE Models. Symposium on Modeling of Complex Process, Texas.

## **Capítulo 5**

### **Otimização Dinâmica**

Os desenvolvimentos relatados nos dois últimos capítulos foram aplicados à otimização dinâmica da transição de grades de um reator de polimerização de polipropileno. Um modelo fenomenológico do reator, previamente desenvolvido, foi utilizado como modelo de planta e como modelo original para ser convertido na versão simplificada, base da otimização, como mostrado no Capítulo 4.

#### **5.1 Proposta de um Otimizador Em Tempo Real**

A implementação do otimizador em tempo real utilizou o algoritmo desenvolvido no Capítulo 3 associado ao método de obtenção de modelos descrito no Capítulo 4. O objetivo desta implementação foi o desenvolvimento de um algoritmo de otimização de processos em tempo real capaz de gerar soluções do problema de otimização extremamente rápidas, ao ponto de ser aplicado não só como otimizador de processos, mas também como controlador. Esta abordagem refere-se àquela descrita na revisão bibliográfica denominada de monocamada.

O otimizador implementado foi desenvolvido em Matlab 5.3 e aplicado ao modelo fenomenológico do processo para simular a planta real. Nesta seção serão descritos os detalhes da implementação que levou a geração dos resultados obtidos.

### 5.1.1 | **Visão Geral do Otimizador**

O otimizador de processos desenvolvido neste trabalho foi dividido em 2 etapas de inicialização e 5 etapas que ocorrem em seqüência, repetida e continuamente, realizando a otimização em tempo real a cada intervalo de amostragem. As etapas de inicialização são divididas em:

- Inicialização da simulação do processo.
  
- Inicialização do otimizador propriamente dito.

A inicialização da simulação do processo consiste em realizar todas as operações necessárias para a simulação do processo industrial, as quais precisam ser feitas uma única vez, no início da simulação. Nesta etapa inicia-se a contagem do número de *scans*<sup>1</sup> do otimizador e calcula-se o estado estacionário do processo ou atribui-se o valor inicial dos estados especificados pelo usuário, quando for o caso. Também nesta etapa são iniciados os valores de horário atual no servidor de dados da planta simulada, bem como os valores das ações de controle e distúrbios para que possam ser lidos pelo otimizador. A última operação desta etapa consiste em avaliar o valor das variáveis de saída para os estados estacionários calculados, de forma que seja possível também disponibilizá-los no servidor de dados.

A inicialização do otimizador ocorre imediatamente após a inicialização da simulação do processo, uma vez que é necessário que se tenha valores válidos das leituras das variáveis de processo para que se realize esta etapa. As operações realizadas na inicialização do otimizador consistem em realizar as leituras das ações de controle iniciais (iguais aos valores medidos destas variáveis), iniciar o valor dos estados do modelo interno do otimizador e determinar o valor das variáveis de saída para os estados iniciais estimados. Nesta etapa também se faz a leitura dos valores iniciais de parâmetros de configuração e operação.

Devido ao fato do MatLab 5.3 não permitir o processamento de múltiplas tarefas, os algoritmos de simulação e otimização foram arrançados de forma a rodarem de forma seqüencial, embora na prática, o otimizador e o processo

---

<sup>1</sup> Scan é o número de execuções de um ciclo completo de otimização que gera uma ação no processo industrial.

industrial gerem valores e resultados de forma concomitante. Assim, realiza-se a simulação do processo para um intervalo de amostragem e posteriormente o cálculo da otimização em tempo real, tomando-se o cuidado de, nesta segunda etapa, utilizar-se os valores medidos das variáveis no instante anterior (penúltimos valores) e não os valores recentemente calculados. Isso garante que o otimizador comporte-se como se estivesse sendo executado em paralelo com a simulação do processo industrial. Para garantir que esta sincronia seja realizada, nenhuma informação é trocada entre o processo e o otimizador entre estas duas etapas de cálculo, o que significa que a leitura de dados do processo para o otimizador é realizada antes da etapa de simulação do processo. No primeiro instante de amostragem esta leitura é realizada pela função de inicialização da otimização e nos demais instantes a leitura dos dados é realizada no final do *loop* anterior.

Após estas duas etapas de cálculo, a ação de controle calculada é escrita no servidor de dados do processo para que a mesma seja considerada no próximo intervalo de integração. Logo após a escrita (etapa 3) e antes que se faça a leitura das variáveis do processo (etapa 5), é necessário realizar um cálculo instantâneo do valor das variáveis de saída do processo. Isso é feito para que se, eventualmente, houver alguma variável de saída que sofra uma influência direta de uma ação de controle, ou que apresente uma dinâmica muito rápida, este efeito seja considerado, constituindo assim a etapa 4 do *loop* de atuação do otimizador sobre o processo.

### **5.1.2 | A Simulação do Processo**

A simulação do processo industrial utilizou o modelo fenomenológico apresentado no Anexo 1 e o solver ode45 do MatLab 5,3. Este solver resolve o sistema de equações diferenciais pelo método de runge-kutta de 4ª ordem refinando a solução com a aproximação de 5ª ordem. A simulação é realizada a partir de um tempo inicial igual ao tempo final da simulação do *loop* anterior e integrada por um período de amostragem após o tempo inicial.

### **5.1.3 | O Otimizador**

O otimizador é composto por duas sub-rotinas:

- Rotina de Reconciliação de dados e estimação de parâmetros;
- Rotina de Otimização Dinâmica.

A rotina de reconciliação de dados e estimação de parâmetros é a responsável por fornecer o *feedback* não-linear do processo para o algoritmo de otimização dinâmica. Para que isso seja feito, é necessário que o modelo do otimizador tenha parâmetros estimáveis (modificáveis) e suficientemente sensíveis para provocar o reposicionamento do modelo dinâmico a fim de torná-lo uma representação mais fidedigna do processo. O método utilizado para realizar esta correção dos parâmetros foi o CEKF (*Constrained Extended Kalman Filter*) cuja descrição será mostrada na seção seguinte.

A Rotina de otimização dinâmica, por sua vez, utiliza o algoritmo GA-GRG mostrado no Capítulo 3 associado em um método seqüencial com os modelos descritos no Capítulo 4. Devido ao fato deste método não necessitar uma integração numérica, o mesmo torna-se extremamente rápido na sua avaliação, eliminando a necessidade de um método de discretização. Os graus de liberdade usados na otimização são as ações de controle de um dado horizonte, sendo que somente o primeiro conjunto de ações é efetivamente aplicado ao processo industrial, tal como no controle preditivo baseado em modelos (MPC).

#### 5.1.4 A Função Objetivo

O otimizador desenvolvido nesta tese utiliza uma formulação denominada “Perseguição de objetivos” (*Goal Attainment*). Esta formulação considera as funções objetivo do problema original como um conjunto de restrições adicionais de um problema transformado, conforme mostrado no Capítulo 3.

As funções objetivo utilizadas em controladores preditivos e em otimizadores em tempo real costumam ser constituídas de uma série de termos que quantificam contribuições de diferentes aspectos importantes do processo para a tomada de uma decisão. Tais termos costumam ser considerados na forma de um somatório em uma função objetivo única (método da soma ponderada), a exemplo do controlador preditivo não-linear BRNMPC (Duraiski et al. 2001) que utiliza uma função objetivo dada pela equação 5.1:

$$J = \min_{\delta U_{[0]_s}^M} \left( \left\| \Gamma_{[0]} \cdot (Y_{[0]} - R_{[0]}) \right\|^2 + \left\| \Lambda_{[0]} \cdot \left( T \cdot (\delta U_{[-1]} + U_{[-2]}^B) \right) \right\|^2 + \left\| \Psi_{[0]} \cdot \left( U_{[-1]}^B + \delta U_{[0]} - Z_{[0]} \right) \right\|^2 + \left\| \Phi \cdot s \right\|^2 \right) \quad (5.1)$$

Nesta representação observa-se a existência de quatro termos distintos. O primeiro termo quantifica o erro da predição do modelo em relação ao *setpoint* especificado; o segundo termo pondera as variações entre ações de controle consecutivas; o terceiro termo pondera o desvio da variável manipulada em relação a um *target* especificado; e o quarto termo pondera a violação das restrições das variáveis controladas. Cada um dos termos descritos é ponderado por um peso que multiplica o valor do termo em meio a um somatório único que constitui a função objetivo.

A proposta desta tese é avaliar os termos importantes da função objetivo de uma forma multiobjetivo MINMAX. Nesta formulação, consideram-se múltiplas funções objetivo e desta forma cada termo considerado importante na otimização é ponderado de uma maneira bastante peculiar sem perder a relação com os demais. Desta forma a função objetivo da equação 5.2:

$$\min_x (w_1 f_1(x) + w_2 f_2(x) + w_3 f_3(x) + \dots + w_n f_n(x)) \quad (5.2)$$

é transformada no problema dado pelas equações 5.3 e 5.4,

$$\min_{\gamma, x} (\gamma) \quad (5.3)$$

sujeito à:

$$F(x) - \gamma \cdot W \leq F_0 \quad (5.4)$$

onde  $F(x)$  é a função vetorial de funções  $f_1(x)$  à  $f_n(x)$ ,  $W$  é um vetor de pesos  $w_1$  à  $w_n$ ,  $F_0$  é um vetor de tolerâncias, e  $\gamma$  é uma variável auxiliar introduzida no problema.

Nesta nova formulação do problema, os termos da função objetivo passam a ser considerados como restrições do problema de otimização e isto implica em uma mudança na concepção do problema. Em ambas as formas da otimização multiobjetivo, se existe conflito entre as funções objetivo, a solução do problema original não é única. Conseqüentemente, a solução obtida por qualquer um dos métodos mencionados vai estar associada aos pesos impostos em cada alternativa. Para tratar da não unicidade de solução é definido o conceito de não inferioridade, também conhecido como otimalidade

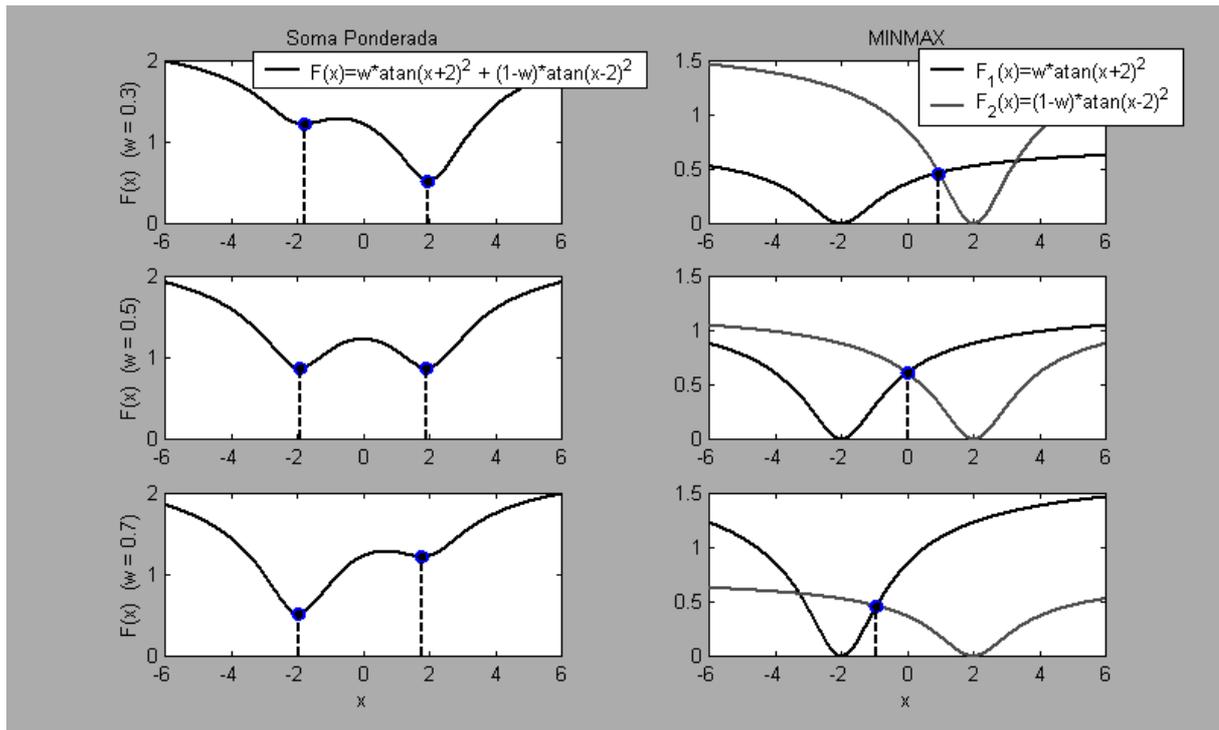
de Pareto. Uma solução não inferior é aquela que uma posterior melhoria em uma das funções objetivo causa a degradação de outras.

A otimização na forma MINMAX tem uma aplicação importante para evitar a não convexidade dos problemas. Quando se utiliza a soma ponderada, a consideração de termos originalmente convexos pode resultar numa combinação não convexa. Para exemplificar esta situação considere os termos mostrados nas equações 5.5 e 5.6:

$$f_1 = \arctan(s + 2)^2 \quad (5.5)$$

$$f_2 = \arctan(s - 2)^2 \quad (5.6)$$

A Figura 5.1 mostra a situação onde os dois termos mencionados anteriormente são otimizados utilizando-se as duas formas de otimização multi-objetivo citadas. Observa-se que na soma ponderada ocorre a formação de uma função objetivo não convexa onde dois ótimos locais coexistem e por mais que se modifique os pesos relativos entre estes termos os ótimos continuam praticamente nos mesmos pontos. Por outro lado, quando estes termos são considerados na forma MINMAX, o problema apresenta um ótimo global e a modificação dos pesos provoca um deslocamento real do ótimo global entre os dois ótimos de cada termo mostrando que esta forma de considerar o problema é mais adequada neste caso.



**Figura 5.1:** Comparação dos ótimos obtidos pelas duas diferentes formas de otimização multiobjetivo para funções não convexas proveniente de termos convexas

Na prática, a aplicação de otimização multiobjetivo em controle avançado ou otimização em tempo real utiliza um ajuste de pesos constantes e, ao longo da operação, as condições às quais o processo é submetido provoca alterações diretamente nos valores dos termos da função objetivo, o que faz com que se mostre uma análise diferente da realizada com o conjunto de Pareto. Na forma da soma ponderada dada pela equação 5.2, cada termo colabora para a penalização da função objetivo de forma proporcional a sua grandeza sem que nunca a participação de algum termo perca totalmente a sua importância. Na forma da perseguição dos objetivos, mostrada nas equações 5.3 e 5.4 eventualmente um dos termos pode perder a sua importância e ser totalmente ignorado desde que o seu valor seja superado por outro termo da função objetivo. Na prática, o novo problema estabelece a condição de que somente o termo de maior valor será minimizado a menos que, na direção de otimização outro termo passe a ter maior valor. Neste caso, o termo anterior é desprezado e o novo termo passa a ter importância.

A Figura 5.2 mostra uma comparação entre o ótimo obtido com os dois métodos para o problema que consiste em encontrar o ótimo de dois termos significativos dados pelas equações 5.7 e 5.8.

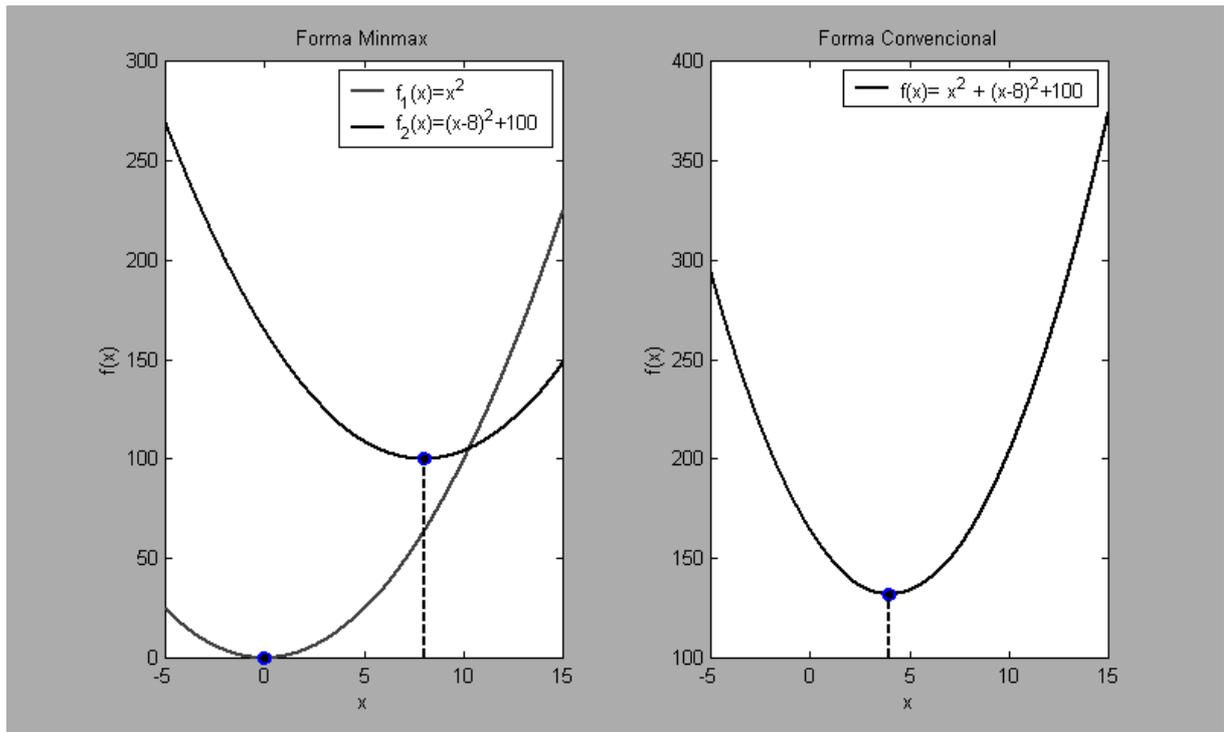


Figura 5.2: Comparação entre os ótimos obtidos pelo método convencional e minmax (caso 1).

$$f_1(x) = x^2 \quad (5.7)$$

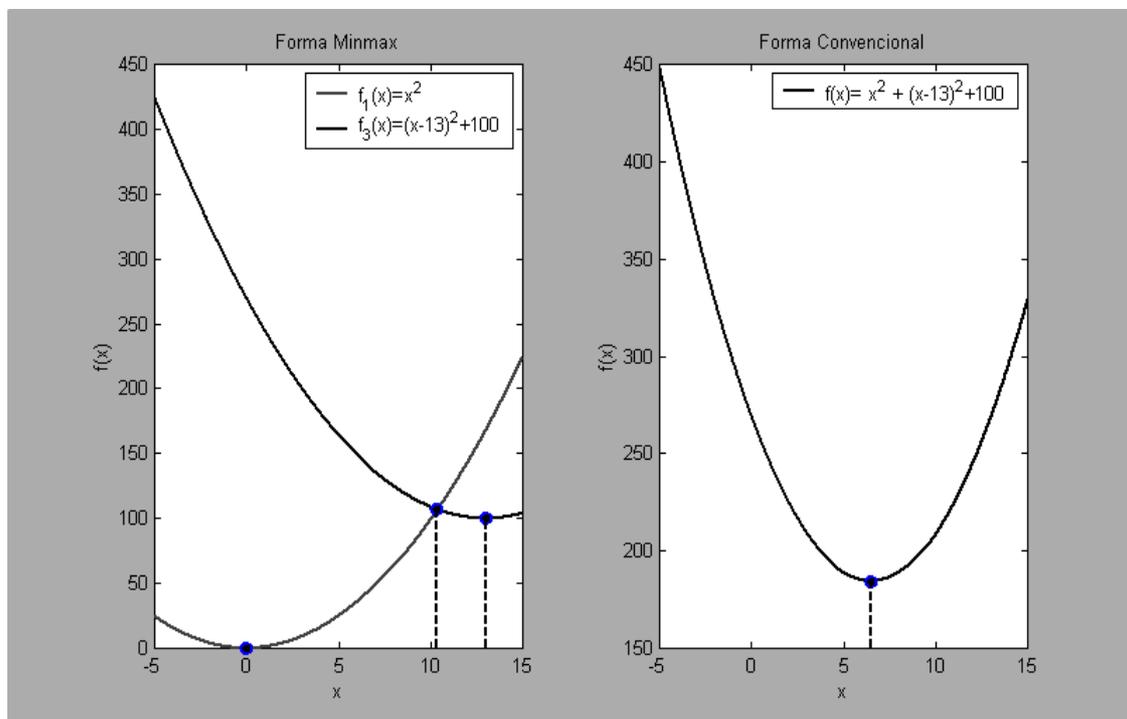
$$f_2(x) = (x-8)^2 - 100 \quad (5.8)$$

Na forma da soma ponderada, mostrada do lado direito da figura, os dois termos são somados e um ótimo é obtido em um valor de  $x = 4$  enquanto que no método minmax cada função apresenta um ótimo ( $x = 0$  para o termo  $f_1$  e  $x = 8$  para o termo  $f_2$ ) sendo que efetivamente o valor de ótimo considerado pelo problema é o ponto  $x = 8$ . Isso ocorre porque o ponto  $x = 8$  é o ótimo de um dos termos considerados e todos os demais valores dos outros termos do problema são menores que o valor do termo considerado otimizado neste ponto. Esta é a principal diferença entre as duas formas de otimização considerando-se pesos constantes.

Apesar das considerações anteriores, nem sempre o ótimo ocorrerá exatamente sobre o ótimo de um dos termos. A Figura 5.3 mostra um caso onde um compromisso entre os dois termos é estabelecido. Neste caso, o termo  $f_2$  é trocado por um termo  $f_3$  que possui o seu ótimo no ponto  $x = 13$ . Observe que neste ponto o valor do termo  $f_1$  possui um valor maior que o do termo  $f_3$  não caracterizando assim o ótimo da função objetivo. Assim, um compromisso surge entre os dois termos e o ótimo passa a ser uma

ponderação de ambos. Note também que, embora o compromisso exista, o ponto ótimo obtido pelas duas formas de propor o problema é diferente ( $x=10,3462$  para o Minmax e  $x=6,5$  para a forma da soma ponderada) o que mostra que os pesos entre as diferentes formas não tem uma relação direta, já que nestes exemplos utilizou-se peso unitário para todos os termos.

É importante ressaltar que nas duas situações das Figuras 5.2 e 5.3 o ótimo do problema pode ocorrer em qualquer ponto entre os ótimos dos dois termos independentes variando-se apenas a relação entre os pesos, já que esta é a região não inferior do problema, e que as análises anteriores foram feitas com o intuito de mostrar o efeito da utilização de pesos constantes.



**Figura 5.3:** Comparação entre os ótimos obtidos pelo método convencional e minmax (caso 2)

No caso de controle preditivo, esta consideração torna-se importante quando se analisa as contribuições dos termos de erro e supressão de movimentos, representados respectivamente pelos termos 1 e 2 da função objetivo mostrada na equação 5.1. Quando a otimização ocorre da forma convencional (soma ponderada), o termo de supressão de movimentos pondera toda e qualquer ação de controle que desloque a variável manipulada do seu valor atual, independentemente da distância que o sistema encontra-se do *setpoint*. Se o mesmo problema for considerado na forma *minmax*, a ponderação da ação de controle ocorrerá de formas diferentes para mudanças

maiores ou menores de setpoint. Nas pequenas mudanças que deslocarem pouco o valor da variável manipulada, o termo da supressão de movimentos terá um valor muito pequeno frente ao termo do erro de forma que somente o termo do erro contribuirá para a determinação do valor otimizado da função objetivo. Por outro lado, as grandes mudanças que necessitem um deslocamento maior da variável manipulada levarão o termo referente à supressão de movimentos a possuir um valor maior e um compromisso entre os dois termos pode vir a surgir. Isso leva o controlador a considerar a supressão de movimentos apenas para os casos em que os grandes deslocamentos da variável manipulada forem necessários.

Cabe salientar que todas as implementações de controladores preditivos são realizados usando a versão ponderada, pois esta formulação é facilmente convertida em um problema de programação quadrática, o qual é facilmente resolvido. Entretanto, como discutido aqui, este comportamento diferenciado apresentado pela formulação *minmax* pode apresentar uma série de vantagens que vão desde o ajuste até o desempenho e robustez apresentado pelo controlador. Estas características, embora não estudadas nesta tese, devem ser futuramente estudadas em detalhes.

### **5.1.5 O Filtro de Kalman Estendido**

Os filtros de Kalman são algoritmos matemáticos capazes de proporcionar a reconciliação dinâmica do processo e promover a estimação de parâmetros de modelos que os descrevem. A metodologia fundamentada na década de 60 passou por uma série de evoluções ao longo dos anos, sendo as mais importantes e significativas a geração dos filtros de Kalman Estendido (EKF) e o Filtro de Kalman Estendido com Restrições (CEKF).

No presente otimizador de processos, optou-se pela utilização do CEKF (Salau. et al. 2007) pelo fato deste algoritmo trabalhar com restrições operacionais, as quais permitem que o valor dos estados e parâmetros sejam limitados a valores fisicamente palpáveis, sem que haja a estimação ou reconciliação de valores sem significados físicos. São muitas as variações nas implementações destes algoritmos, e sendo assim, a implementação efetivamente utilizada está sendo apresentada nesta seção.

O algoritmo consiste em duas etapas básicas denominadas:

- predição;
- correção.

A etapa da predição é realizada para que, com base no valor das variáveis de entrada do processo, realize-se a estimativa do que está acontecendo no processo com os estados e saídas no momento atual. Esta etapa consiste basicamente na integração numérica do modelo diferencial que representa o processo, utilizando-se os valores de variáveis de entrada medidos no instante de amostragem anterior e o valor dos estados iniciais condizentes com os valores corrigidos pela reconciliação, também no instante de amostragem anterior.

A etapa de correção, por sua vez, consiste na otimização dos valores dos estados, parâmetros e saídas, de forma que se obtenham os menores erros possíveis entre os valores dos estados e parâmetros e suas predições dadas pelo modelo, bem como a menor diferença possível entre o valor das saídas e seus valores medidos do processo. Como modelo e processo apresentam divergências moderadas entre seus valores estimados, surge um compromisso entre a obtenção do menor erro em relação ao modelo e o menor erro em relação ao processo, que, quando devidamente ponderados, permitem que se encontre em valor ótimo (mínimo) de erro entre o estado corrigido e o valor de estado real do processo.

No CEKF implementado nesta tese, estas duas etapas podem ocorrer de forma seqüencial, nos casos em que se tem apenas reconciliação, ou de forma simultânea, para os casos que se tenha estimativa de parâmetros, como mostram os passos do algoritmo detalhados como segue.

O primeiro passo deste algoritmo consiste em realizar uma pré-predição do estado atual. Este nome é dado a este passo devido ao fato de que a predição realizada neste passo não necessariamente será a predição efetivamente utilizada na etapa de correção do algoritmo. Isso se deve ao fato de que, se parâmetros estiverem sendo estimados a predição tem que ser corrigida durante a otimização.

O segundo passo do algoritmo consiste em realizar a predição e correção da matriz de covariância dos estados do modelo. Assim como os

estados são preditos e corrigidos, é necessário fazer com que a matriz de covariância dos estados passe pelo mesmo processo, uma vez que esta matriz indica o grau de incerteza que se tem entre os estados calculados e seus valores reais. Isso significa que sempre que o valor do estado é recalculado, seja pela predição ou correção, a incerteza que se tem nestes valores também é modificada, o que faz com que seja necessário recalculá-la a cada nova estimativa de estados. Assim sendo, existem etapas de predição e correção da matriz  $P$ , que especificamente neste algoritmo, são realizadas em um único cálculo. Ao contrário dos estados, a formulação da atualização da matriz  $P$  sugere que esta seja primeiramente corrigida no instante anterior e posteriormente predita para o instante atual como segue (Salau. et al. 2007):

$$P_{corr} = Q + \Phi \cdot P \cdot \Phi^T - \Phi \cdot P \cdot H^T \cdot (H \cdot P \cdot H^T + R)^{-1} \cdot H \cdot P \cdot \Phi^T \quad (5.9)$$

onde:  $P_{corr}$  é a matriz de covariância dos estados corrigida no instante de tempo anterior e predita para o instante de tempo atual;  $Q$  é a matriz de incerteza dos estados;  $H$  é a jacobiana das equações que relacionam os estados com as variáveis de saída;  $\Phi$  é a versão discretizada da Jacobiana do modelo em relação aos estados, ou seja:

$$\Phi = e^{(F \cdot \Delta t)} \quad (5.10)$$

onde  $\Delta t$  é o tempo de amostragem e  $F$  é a matriz jacobiana dada por:

$$F = \frac{df}{d \begin{bmatrix} x \\ p \end{bmatrix}} \quad (5.11)$$

onde  $f$  é o sistema de equações que relacionam os estados com as suas derivadas de forma explícita dada por:

$$\frac{dx}{dt} = f(x, p, u) \quad (5.12)$$

É importante destacar que a jacobiana em questão é obtida através da linearização dinâmica do modelo considerando os valores das variáveis no instante de amostragem anterior, pois não só a correção que está sendo realizada nesta formulação refere-se à correção da matriz  $P$  do instante de tempo anterior, como também a predição é a do instante anterior para o atual.

O terceiro passo do algoritmo consiste em realizar a otimização do problema, ou seja, a correção propriamente dita. Para tanto, utiliza a função objetivo que segue:

$$J = \min_{w,v} (w \cdot P^{-1} \cdot w + v \cdot R^{-1} \cdot v) \quad (5.13)$$

Sujeito à:

$$x_{pred} = \text{solução de } \left( \frac{dx}{dt} = f(x, p_{estim}, u_{ant}) \right) \text{ para o intervalo de } t = \Delta t \text{ e condição inicial } x = x_{corr,ant}, t = 0 \quad (5.14)$$

$$w = \begin{bmatrix} x_{corr} - x_{pred} \\ p_{estim} - p_{estim|ant} \end{bmatrix} \quad (5.15)$$

$$v = y_{med} - y_{corr} \quad (5.16)$$

$$y_{corr} = g(x_{corr}, p_{estim}, u_{ant}) \quad (5.17)$$

onde  $w$  é o erro entre o estado predito para o instante de tempo atual e o seu valor corrigido;  $v$  é o erro entre o valor corrigido das variáveis de saída do modelo e os seus valores medidos diretamente do processo;  $P$  é a matriz de covariância dos estados calculada no segundo passo do algoritmo, conforme mostrado anteriormente; e  $R$  é a matriz de covariância das variáveis de saída, ou seja, medidas;  $x_{pred}$  é a predição do estado atual dado pelo modelo;  $x_{corr|ant}$  é o estado corrigido no instante de tempo anterior;  $p_{estim}$  é o valor estimado para o parâmetro;  $u_{ant}$  é o valor das variáveis de entrada no intervalo de amostragem anterior;  $y_{corr}$  é o valor das saídas do modelo avaliado para estado corrigido atual  $x_{corr}$ ;  $y_{med}$  é o valor da variável de saída medido do processo no instante atual;  $g$  é o conjunto de funções que relaciona os estados do modelo às variáveis de saída.

A primeira restrição apresentada para a solução deste problema de otimização trata da resolução do sistema de equações diferenciais que representa o modelo dinâmico do processo. No entanto, somente será necessário resolver este sistema de equações nos casos em que parâmetros estejam sendo estimados e, portanto, influenciando no valor de  $x_{pred}$ . Observe que se não houver variação no valor dos parâmetros durante a otimização, o valor de  $x_{pred}$  é constante e pode ser substituído pelo valor dos estados

previamente preditos fora do problema de otimização no primeiro passo do algoritmo do CEKF.

Nos casos em que a estimação de parâmetros é realizada, uma aproximação linear do problema foi proposta para evitar que o problema de otimização tenha que integrar o sistema de equações diferenciais a cada avaliação das restrições. Esta aproximação consiste em avaliar a influência que uma variação no valor dos parâmetros provoca no valor dos estados preditos. A forma como isso é realizado é dada pela equação que segue:

$$x_{pred} = x_{pre-pred} + \Phi_p \cdot w_p \quad (5.18)$$

onde  $x_{pre-pred}$  são os estados previamente preditos no primeiro passo;  $w_p$  é a parte do vetor  $w$  referente ao erro dos parâmetros; e  $\Phi_p$  são as últimas colunas da matriz  $\Phi$  que relaciona a influência dos parâmetros sobre os estados.

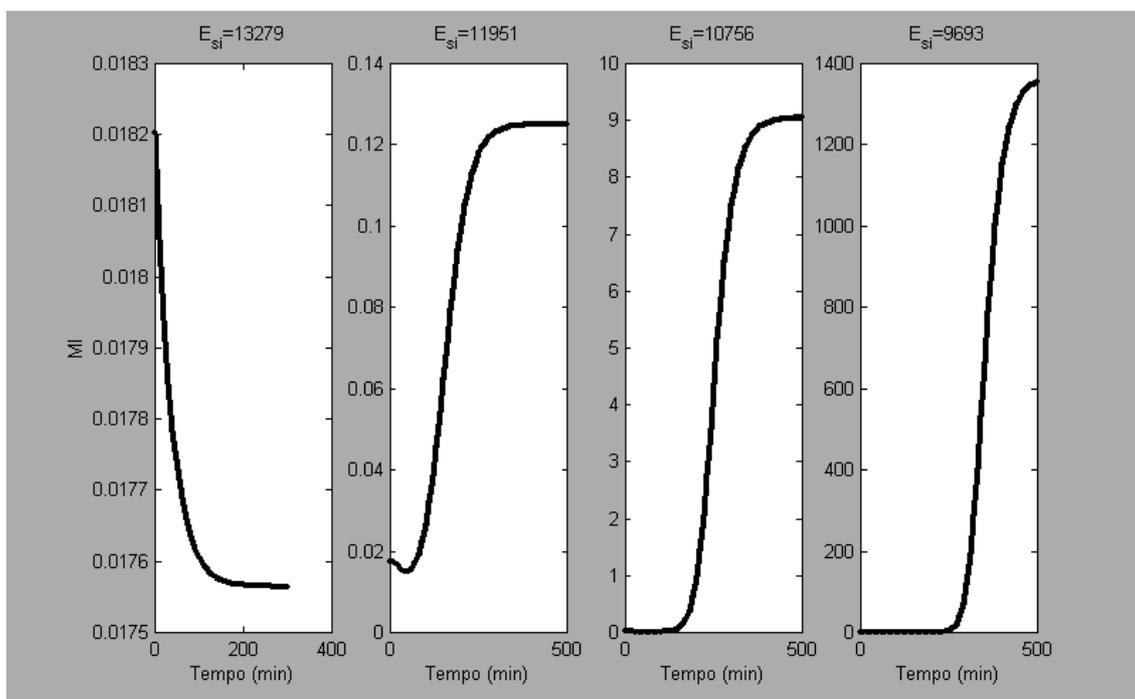
## 5.2 O Modelo do Reator de Polipropileno

O modelo de polimerização usado nesta tese é detalhado no Anexo 1. O processo consiste em converter propeno em polipropileno em solução de hexano. A reação de polimerização ocorre na fase líquida, enquanto concentrações significativas de propeno e hidrogênio encontram-se no domo do reator na fase gasosa. Embora o modelo proposto tenha sido desenvolvido para trabalhar com diferentes catalisadores e múltiplos sítios ativos, os dados cinéticos somente foram obtidos para um catalisador específico. Devido à dificuldade de obter dados cinéticos para outros catalisadores na literatura, foi necessário determinar um conjunto apropriado de dados cinéticos para descrever o comportamento do reator durante a produção de outro grade de polímero.

O procedimento para realizar esta determinação foi modificar o valor de constantes cinéticas do catalisador atual e calcular o estado estacionário do modelo para o novo conjunto de parâmetros cinéticos. Considerando que modificar todas as constantes cinéticas não seria possível, uma vez que um número elevado de reações estão presentes no sistema, a constante escolhida para sofrer as modificações foi a energia de ativação da reação de terminação. A razão para tal escolha é o fato de que o comprimento médio da cadeia polimérica é fortemente dependente da velocidade da reação de terminação.

Se as reações de terminação são dificultadas, o comprimento médio das cadeias poliméricas é aumentado e vice-versa. O primeiro cálculo de estado estacionário do modelo mostrou que o índice de fluidez do polímero gerado no reator utilizando os parâmetros cinéticos originais é igual a 0,95. Considerando que os grades industriais desejados variam entre 1, 10, 100 e 1000, o objetivo desta determinação preliminar foi identificar um catalisador capaz de produzir um polímero com índice de fluidez igual a 1000. Conseqüentemente, a mistura destes dois catalisadores é capaz de produzir polímeros de qualquer um dos grades intermediários.

Uma metodologia mais apropriada para determinação do valor de energia de ativação que leve a produção de um polímero com índice de fluidez igual a 1000 seria a aplicação de um método de resolução de equações algébricas tal como Newton-Rapson, bissecção, entre outros, tendo como variável o valor da energia de ativação da reação de terminação. Considerando-se que não é necessário encontrar um valor exato de energia de ativação que leve o índice de fluidez a 1000, mas sim um valor qualquer acima deste, optou-se por variar manualmente a energia de ativação até que o valor do índice de fluidez no estado estacionário fosse superior a 1000. A figura 5.4 mostra as perturbações realizadas para a determinação da energia de ativação do novo catalisador.



**Figure 5.4:** Perturbações realizadas para determinar as constantes cinéticas do novo catalisador.

Utilizando-se um conjunto pequeno de soluções obteve-se a nova energia de ativação da reação de terminação do novo catalisador.

Para realizar a otimização da transição, escolheu-se as duas variáveis manipuladas mais sensíveis em relação à variável de interesse – Índice de fluidez. São elas: a fração de catalisador A (Original) em relação à alimentação total de catalisador e a vazão de hidrogênio. Como o interesse era apenas a otimização da transição e não a produção, optou-se por otimizar apenas o índice de fluidez do polímero, uma vez que esta é a variável que indica a qualidade e o tipo de produto. Uma segunda modificação foi realizada no modelo original, com o intuito de transformar as variáveis de entrada em grandezas de mesma ordem. Ao invés de se utilizar a vazão de hidrogênio como variável de entrada, utilizou-se a fração de vazão total de hidrogênio estabelecida em 62 g/min. Como não se tinha o valor correto do limite superior para tal vazão arbitrou-se este valor com base nos resultados das simulações que indicavam uma sensibilidade significativa do índice de fluidez em relação à vazão de hidrogênio. Assim a variável manipulada era dada pela fração da alimentação total e ficava limitada no intervalo [0, 1]. Utilizou-se, também, a vazão de catalisador original do modelo ( $F_{cat\_iso} = 4,8$  g/min e  $F_{cat\_atat} = 0,2$  g/min, onde  $F_{cat\_iso}$  é a vazão de catalisador com sítios isotáticos e  $F_{cat\_atat}$  é a vazão de catalisador com sítios atáticos) como sendo a vazão total de catalisador alimentada, sendo a variável de otimização a fração de catalisador A contida nesta vazão. Cabe salientar que considerou-se a quantidade de sítios atáticos e isotáticos constante nos catalisadores A e B.

### 5.3 Conversão do Modelo Fenomenológico

Utilizando-se as duas variáveis de entrada do modelo anteriormente estipuladas, determinou-se uma malha retangular de pontos espalhados pelo espaço bidimensional formado pelas variáveis. A tabela 5.1 mostra os pontos utilizados para a formulação da malha.

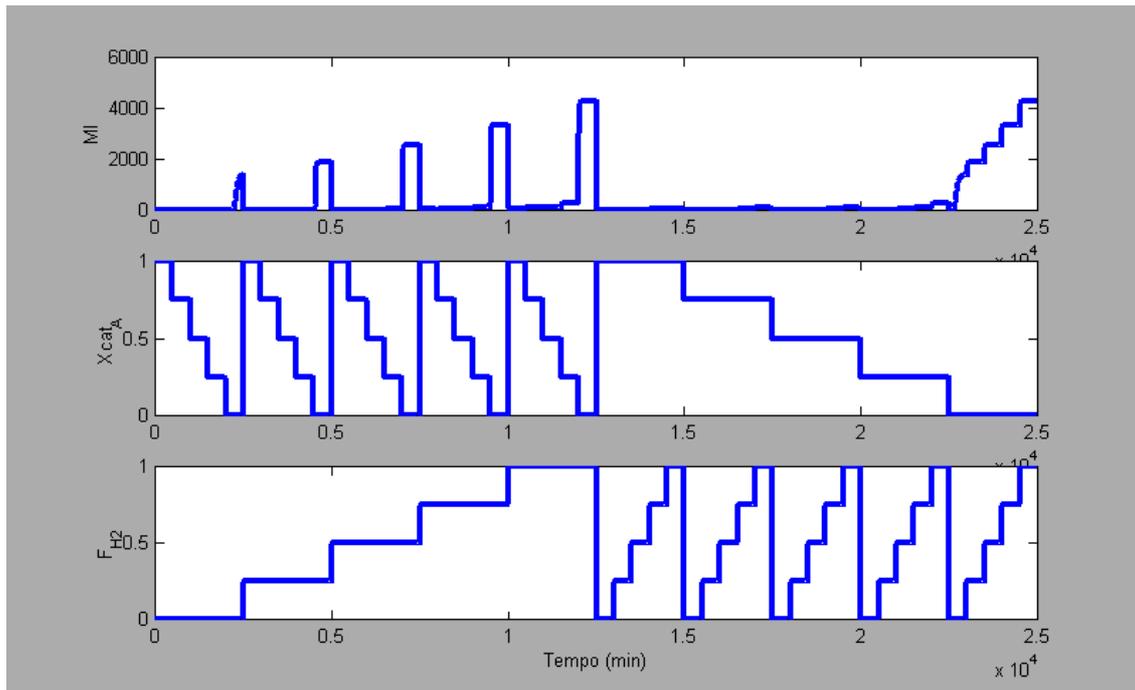
**Tabela 5.1:** Pontos da malha retangular utilizados para a conversão do modelo

Variável	Ponto 1	Ponto 2	Ponto 3	Ponto 4	Ponto 5
Fração de Catalisador	0,00	0,25	0,50	0,75	1,00
Vazão de H <sub>2</sub> normalizada	0,00	0,25	0,50	0,75	1,00

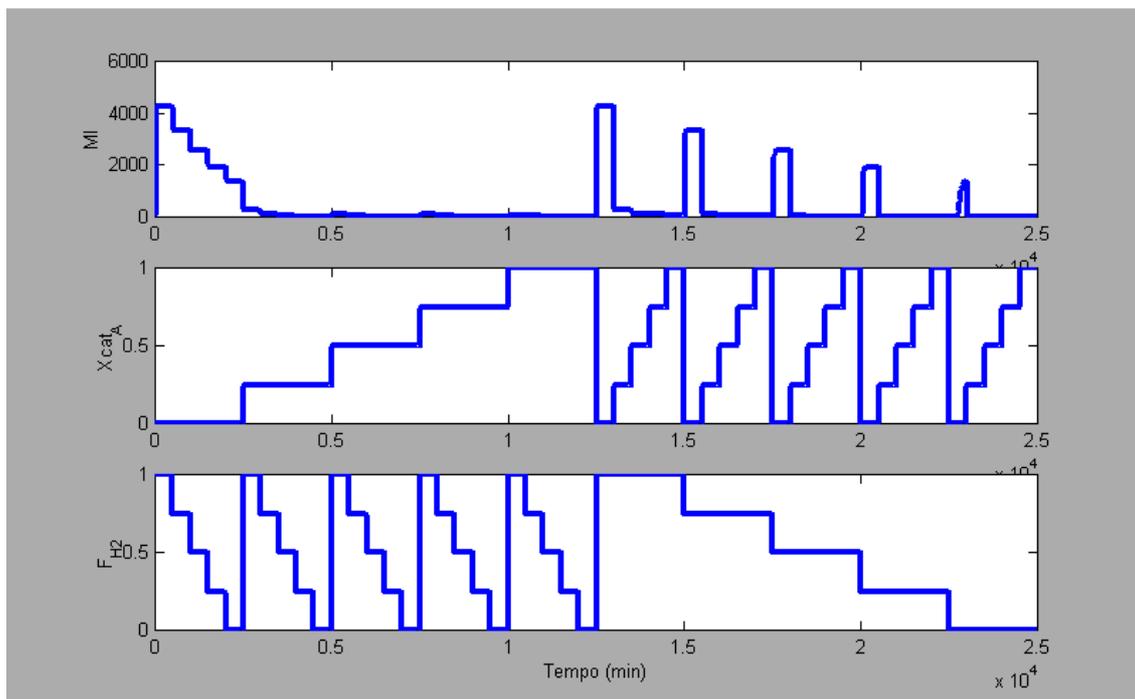
Com a malha retangular de pontos, procedeu-se à simulação das perturbações com o modelo original. As figuras 5.5 e 5.6 mostram o conjunto

de perturbações e as respectivas respostas no índice de fluidez. Nestas figuras, é possível ver que somente quando o catalisador A não é alimentado o índice de fluidez atinge valores elevados, sendo que uma pequena injeção de catalisador A é o suficiente para levar o índice de fluidez a valores baixos. Isso mostra que o  $M_i$  tem um comportamento bastante não linear em função da fração de catalisador alimentada.

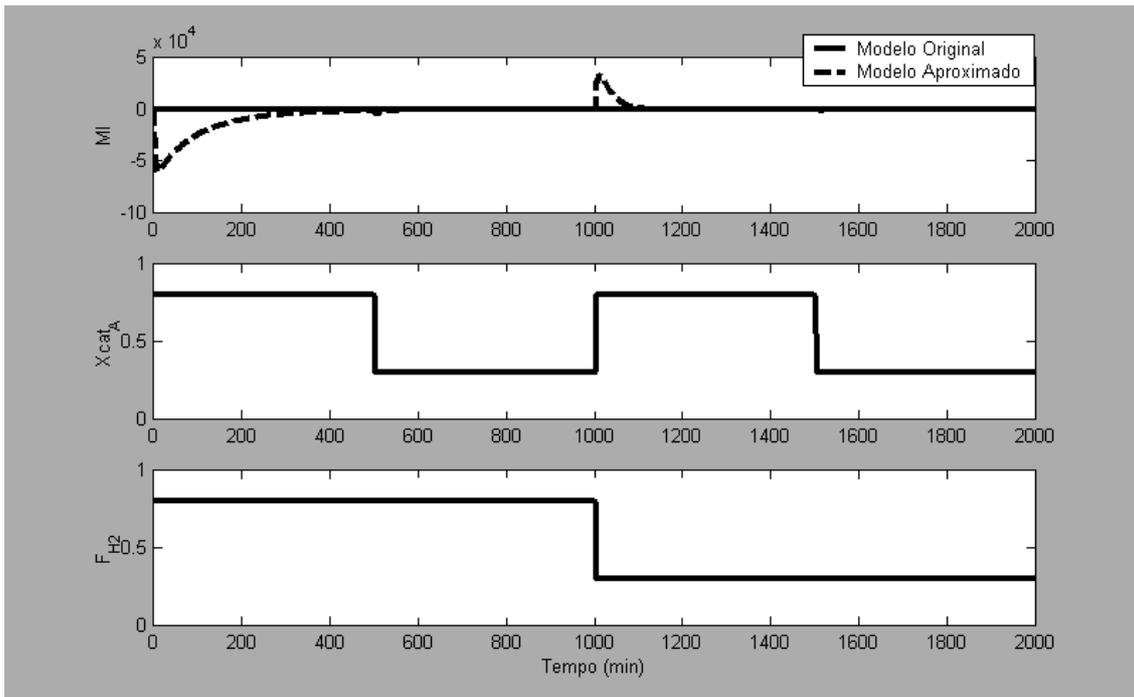
Utilizando-se a técnica mostrada no Capítulo 4, procedeu-se à conversão do modelo original em modelo simplificado, utilizando como entradas a fração de catalisador A ( $X_{cat_A}$ ) e Vazão de hidrogênio normalizada ( $F_{H_2}$ ), e saída o índice de fluidez. Para tanto, se estabeleceu um total de 4 estados no modelo a fim de identificar todas as transições como sendo funções de transferência de quarta ordem. O resultado destas identificações pode ser visto nas figuras 5.7, 5.8 e 5.9. A figura 5.7 mostra o resultado das perturbações (gráficos intermediário e inferior) no MI (gráfico superior). Nesta figura é possível observar o surgimento de dinâmicas (respostas inversas e *overshoot*) não existentes, ou muito maiores do que as existentes no modelo original. A título de comparação, com o modelo original, a figura 5.8 mostra o resultado de cada uma das perturbações mencionadas. A figura 5.9 mostra o resultado das perturbações no modelo aproximado e no modelo original ampliado (multiplicado por 100). Nestas duas figuras, é possível observar que as dinâmicas com respostas inversas e *overshoots* são extremamente ampliadas e, em alguns pontos, criadas sem que existam no modelo original.



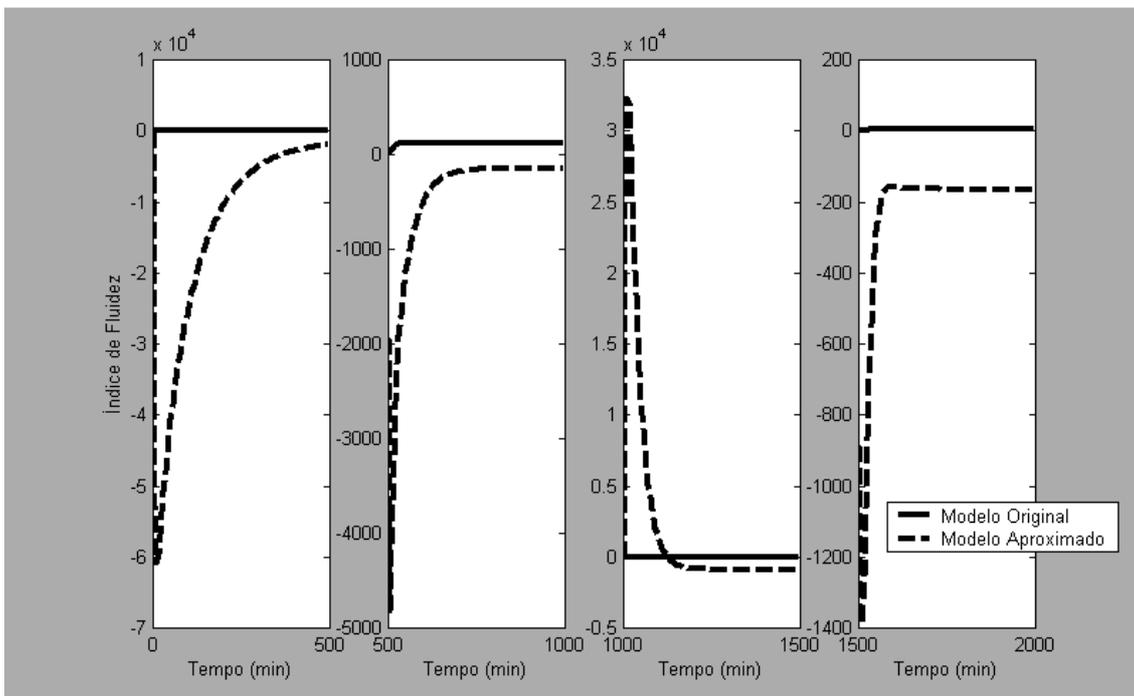
**Figure 5.5:** Perturbações realizadas para realizar a conversão do modelo fenomenológico em simplificado (sentido direto)



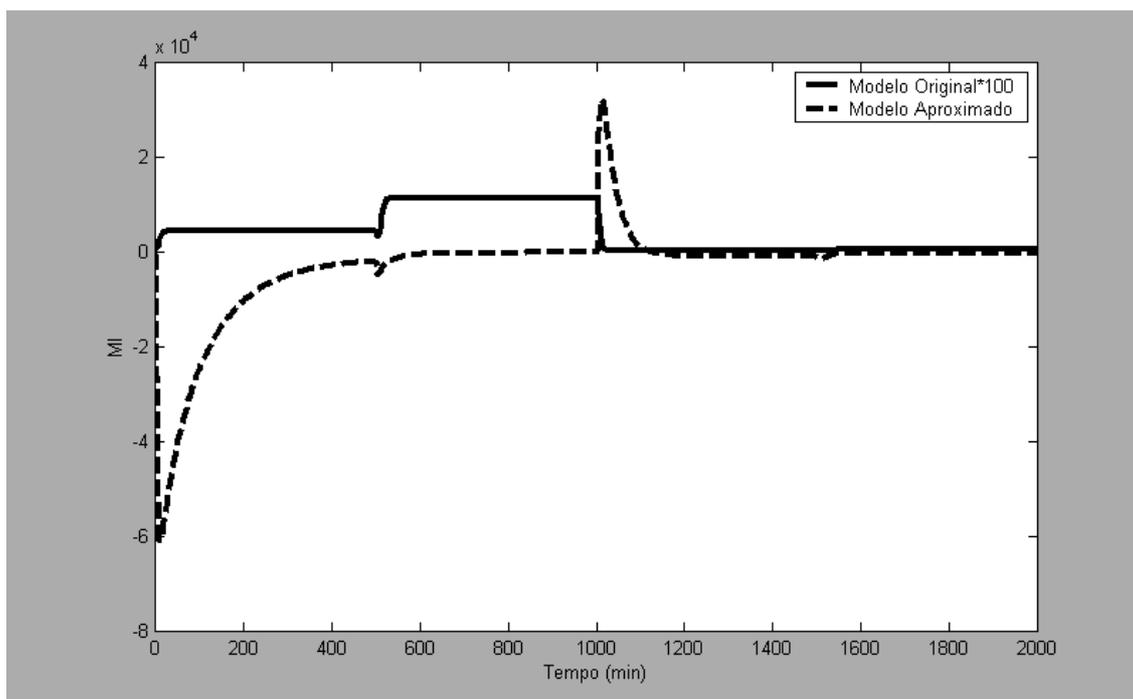
**Figure 5.6:** Perturbações realizadas para realizar a conversão do modelo fenomenológico em simplificado (sentido reverso)



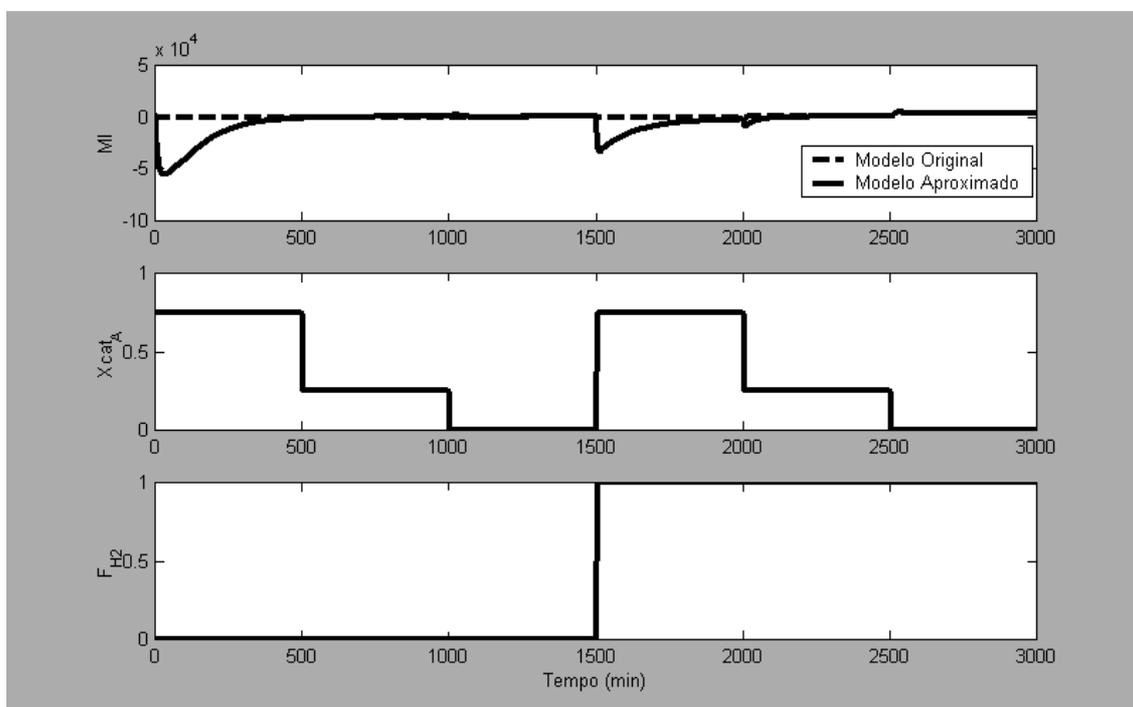
**Figure 5.7:** Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada diferentes das utilizadas na identificação.



**Figure 5.8:** Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada diferentes das utilizadas na identificação. Visualização individual de cada perturbação



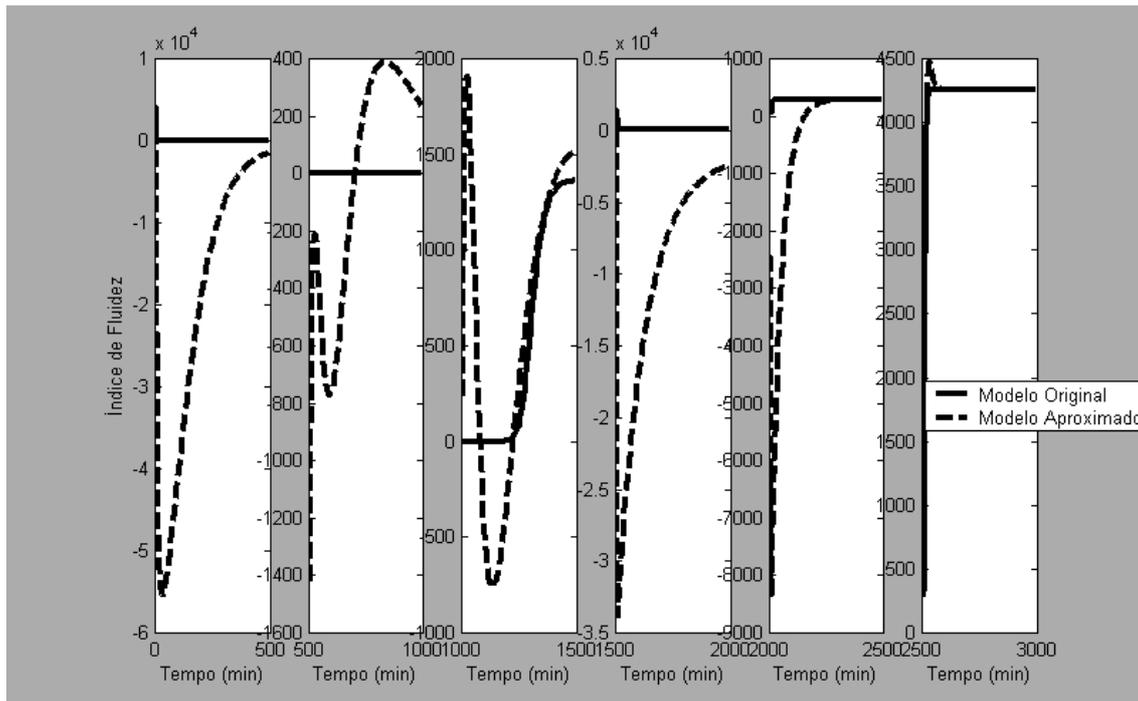
**Figure 5.9:** Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada diferentes das utilizadas na identificação. Amplificação do modelo original.



**Figure 5.10:** Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada iguais às utilizadas na identificação.

A fim de verificar o motivo pelo qual a conversão do modelo apresentou qualidade tão ruim, realizou-se perturbações em ambos os modelos cujos

estados estacionários seriam os mesmos obtidos na identificação do modelo aproximado. As figuras 5.10, 5.11 e 5.12 mostram os resultados destas perturbações. A figura 5.10 mostra as perturbações nas variáveis de entrada juntamente com a resposta no MI. A figura 5.11 mostra cada perturbação individualmente a fim de visualizar as dinâmicas das respostas e a figura 5.12 mostra a comparação da respostas dos dois modelos, sendo o original amplificado por um fator de 10.

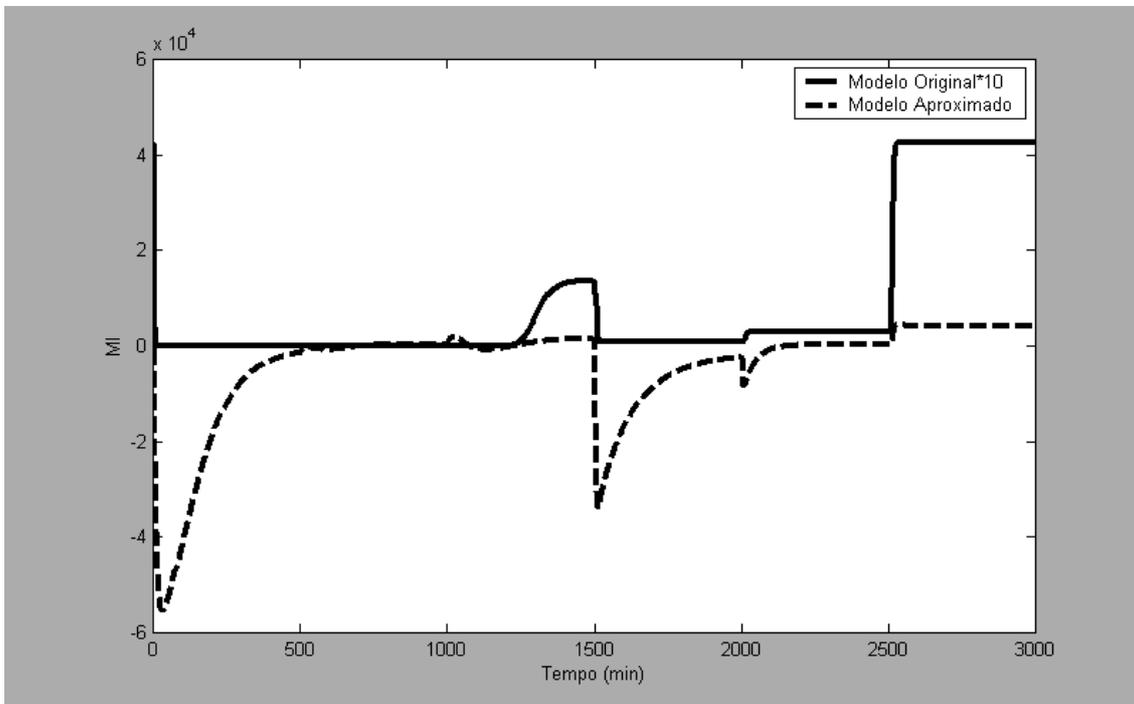


**Figure 5.11:** Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada iguais às utilizadas na identificação. Visualização individual das perturbações

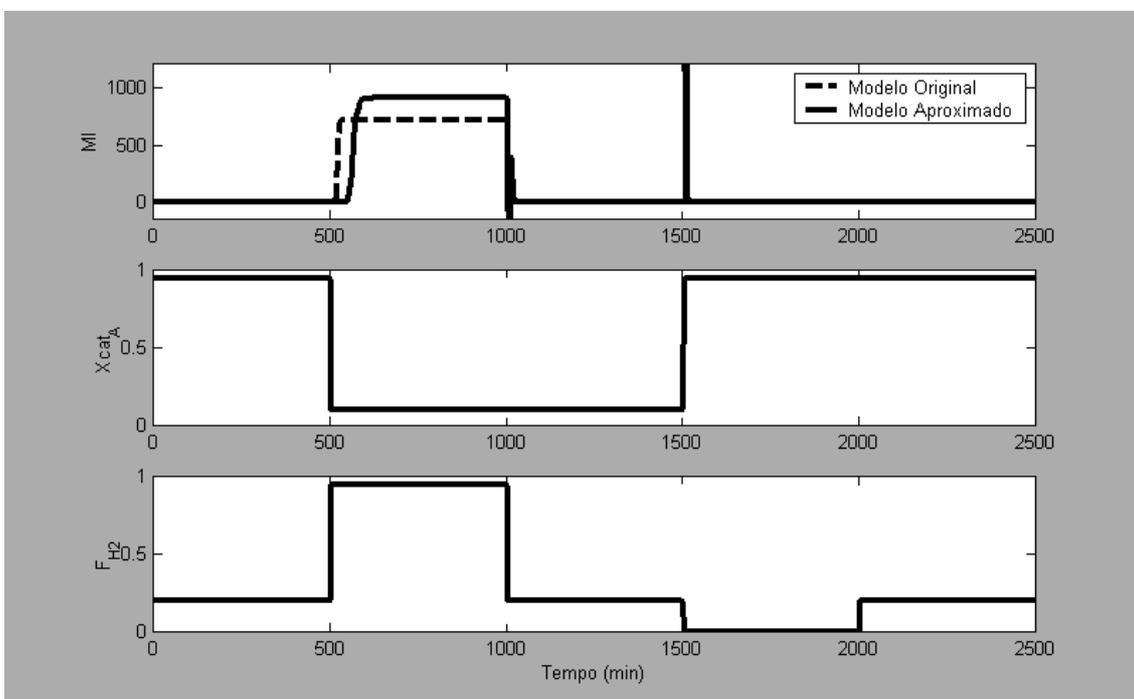
Nas últimas figuras mencionadas, é possível verificar que o valor estacionário do modelo aproximado é condizente com o valor de estado estacionário do modelo original, indicando que os estados estacionários originais são bem descritos pelo modelo aproximado, no entanto a interpolação de estados estacionários realizada pelo modelo aproximado leva a valores que não condizem com o modelo original. Além disso, as dinâmicas geradas no modelo aproximado revelaram-se muito mais pronunciadas do que no modelo original indicando uma péssima qualidade da identificação.

Como foi visto no conjunto de perturbações utilizados para converter o modelo, o índice de fluidez MI apresenta uma grande não linearidade estática o que pode ter sido o causador da má qualidade da conversão, já que a

interpolação dos estados estacionários está sendo realizada por um polinômio de terceira ordem.



**Figure 5.12:** Comparação entre as respostas do modelo original e aproximado de MI frente a perturbações nas variáveis de entrada iguais às utilizadas na identificação. Resposta do modelo original amplificada.



**Figure 5.13:** Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Perturbações diferentes das utilizadas na identificação.

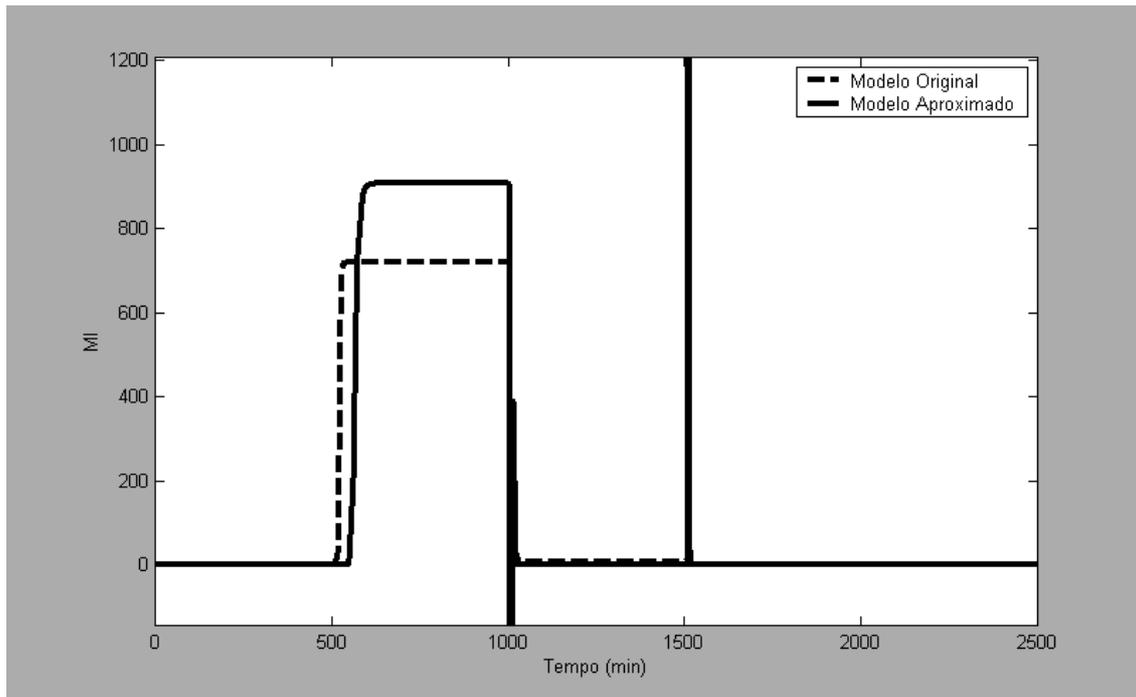
Para resolver o problema da não-linearidade estática do modelo, optou-se por converter modelos cujas saídas fossem os momentos do polímero e não o índice de fluidez diretamente. Tal escolha se deve ao fato de que a maior parte da não linearidade encontrada se deve a forma como o índice de fluidez é calculado a partir dos momentos do polímero. Os momentos dos polímeros estão diretamente relacionados à concentração deste no meio reacional, sendo provenientes de relações de balanço de massa e que, para casos como este onde a temperatura mantém-se constante, obedecem uma relação muito pouco não-linear. Já o índice de fluidez é calculado a partir destes momentos pela equação:

$$MI = \left( \frac{111525}{M_w} \right)^{\left( -\frac{1}{0,288} \right)} \quad (5.19)$$

$$M_w = \left( \frac{mi_2}{mi_1} \right) \quad (5.20)$$

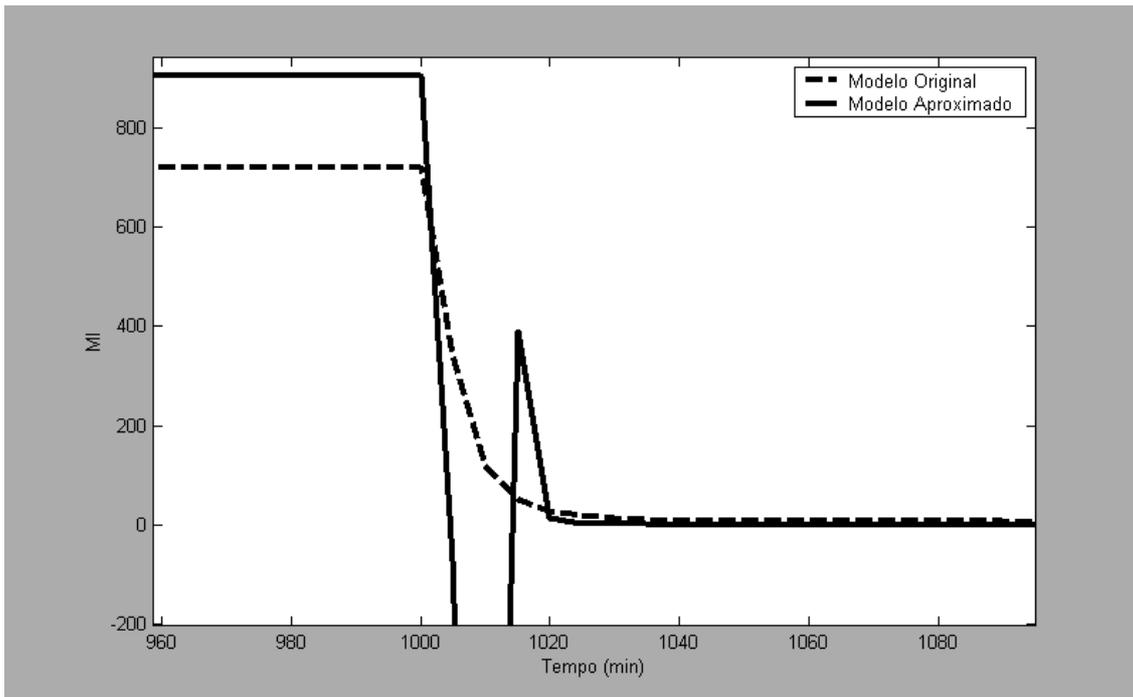
Onde  $mi_1$  e  $mi_2$  são respectivamente os momentos de ordem 1 e 2 do polímero.

Assim sendo, foram identificados modelos de terceira ordem para cada um dos momentos descritos acima, e estas duas novas saídas do modelo aproximado introduzidos na equação 5.19 para que forneçam o valor de MI.

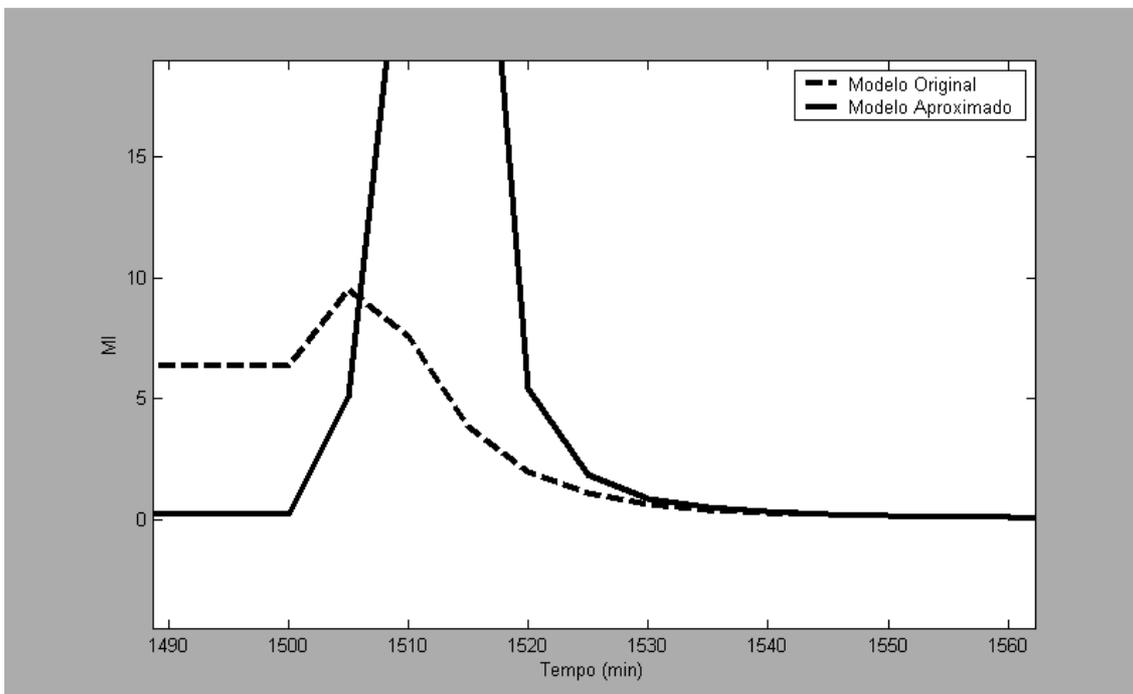


**Figure 5.14:** Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Perturbações diferentes das utilizadas na identificação. Ampliação do gráfico da resposta

A comparação dos modelos obtidos desta forma com o modelo original pode ser vista nas figuras 5.13, 5.14, 5.15 e 5.16. Na figura 5.13, observa-se a resposta dos modelos original e aproximado para um conjunto de perturbações diferentes das utilizadas na conversão do modelo. Em tal figura, observa-se que os estados estacionários do modelo aproximado já são mais condizentes com os estados estacionários preditos pelo modelo original. Na figura 5.14 pode-se observar tal constatação com detalhes. No entanto, a figura 5.15 mostra a terceira perturbação deste conjunto em detalhes onde é possível observar que a dinâmica apresentada pelo modelo aproximado apresenta comportamentos inexistentes no modelo original. Também na figura 5.16 é apresentada a quarta perturbação deste conjunto onde a resposta inversa gerada é muito mais pronunciada do que a existente no modelo original.

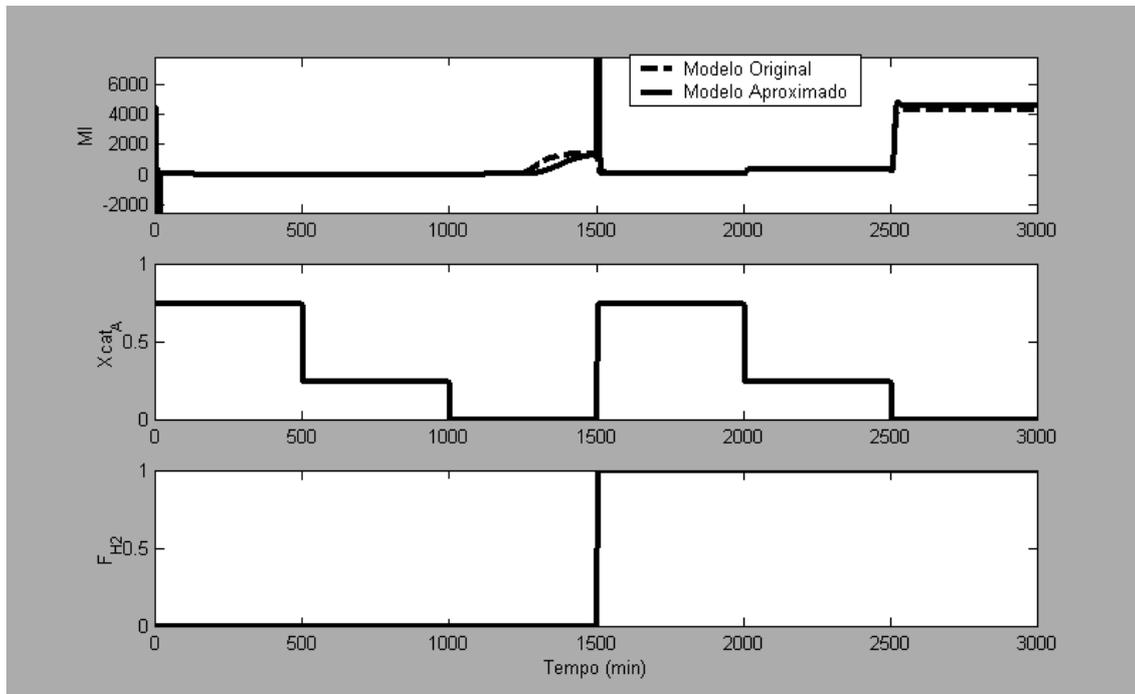


**Figure 5.15:** Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Vazão de catalisador A mantido em 0,2 e Vazão de Hidrogênio modificado de 0,95 para 0,2.

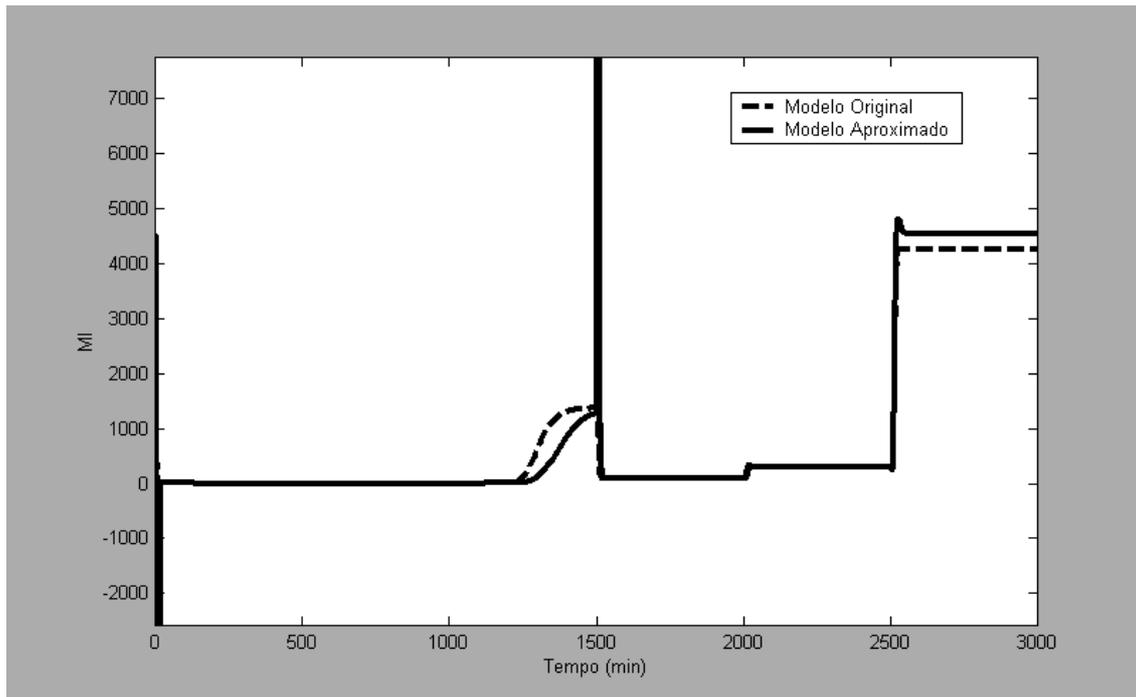


**Figure 5.16:** Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Vazão de catalisador A levado de 0,2 para 0,95 e Vazão de Hidrogênio modificado de 0,20 para 0,00.

Ainda a título de comparação, as Figuras 5.17 e 5.18 mostram a comparação dos modelos original e aproximado frente a perturbações existentes no conjunto de perturbações utilizadas na identificação. Como era de se esperar os estados estacionários neste caso foram rigorosamente descritos pelo modelo aproximado, tal como no modelo original.



**Figure 5.17:** Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações presentes na identificação do modelo aproximado.

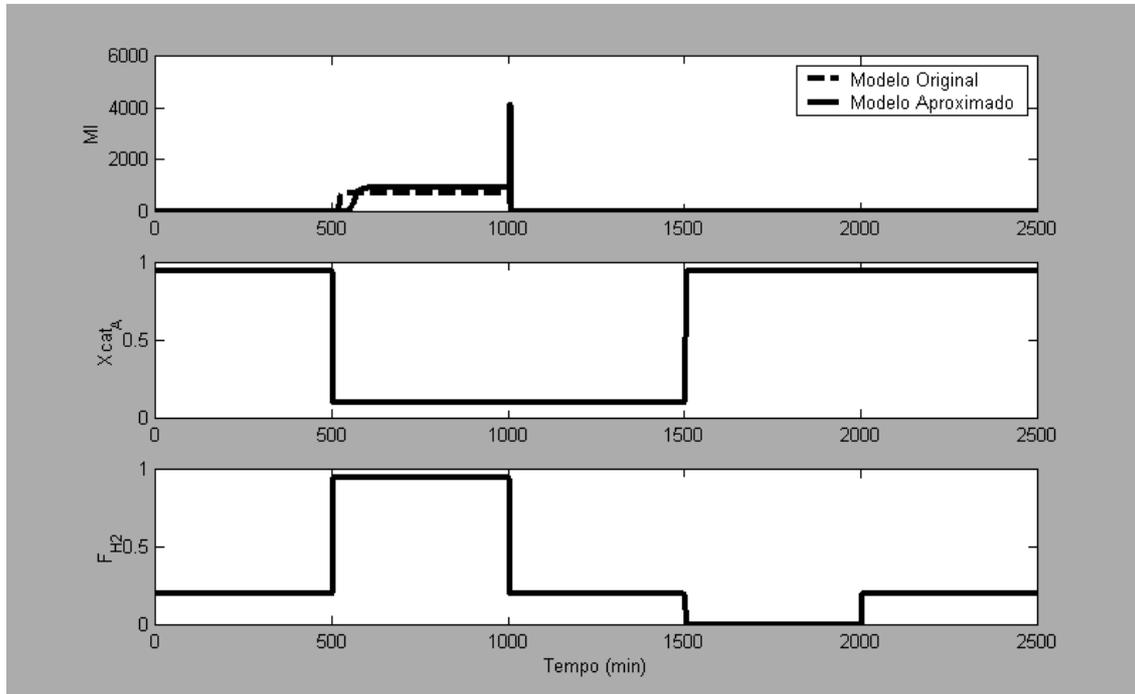


**Figure 5.18:** Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações presentes na identificação do modelo aproximado.

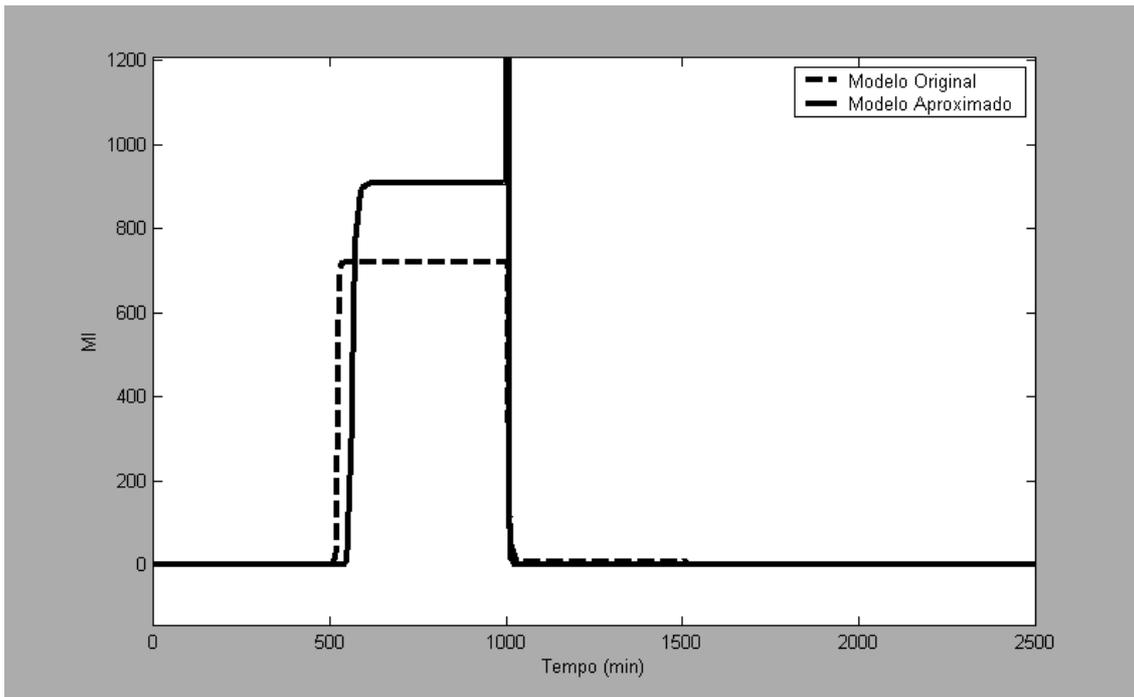
Com o intuito de melhorar a resposta do modelo aproximado, optou-se por reduzir a ordem das identificações realizadas, uma vez que as dinâmicas das respostas identificadas com modelos de terceira ordem apresentavam comportamentos bastante diferentes do modelo original. No entanto, para não perder as manifestações de respostas inversas do modelo original, optou-se por reduzir a ordem dos modelos aproximados para cada momento à ordem 2. A comparação destas novas aproximações podem ser vistas nas Figuras 5.19, 5.20, 5.21 e 5.22. A Figura 5.19 mostra o conjunto de perturbações diferente das utilizadas na identificação do modelo aproximado. Na Figura 5.20, a resposta no MI é mostrada em detalhe onde é possível ver que a qualidade da representação do modelo aproximado mostrou uma certa melhora. No entanto, a Figura 5.21 mostra o detalhe da terceira perturbação onde a descrição aproximada mostra ainda uma resposta inversa inexistente no modelo original. Por outro lado, na figura 5.22 onde a quarta perturbação é mostrada com detalhe, é possível ver que o modelo aproximado ignorou totalmente a presença de uma pequena resposta inversa existente no modelo original.

As Figuras 5.23 e 5.24 mostram a resposta do modelo de segunda ordem baseado em momentos frente a perturbações presentes no conjunto utilizado para a identificação do modelo aproximado. Nestas figuras, é possível

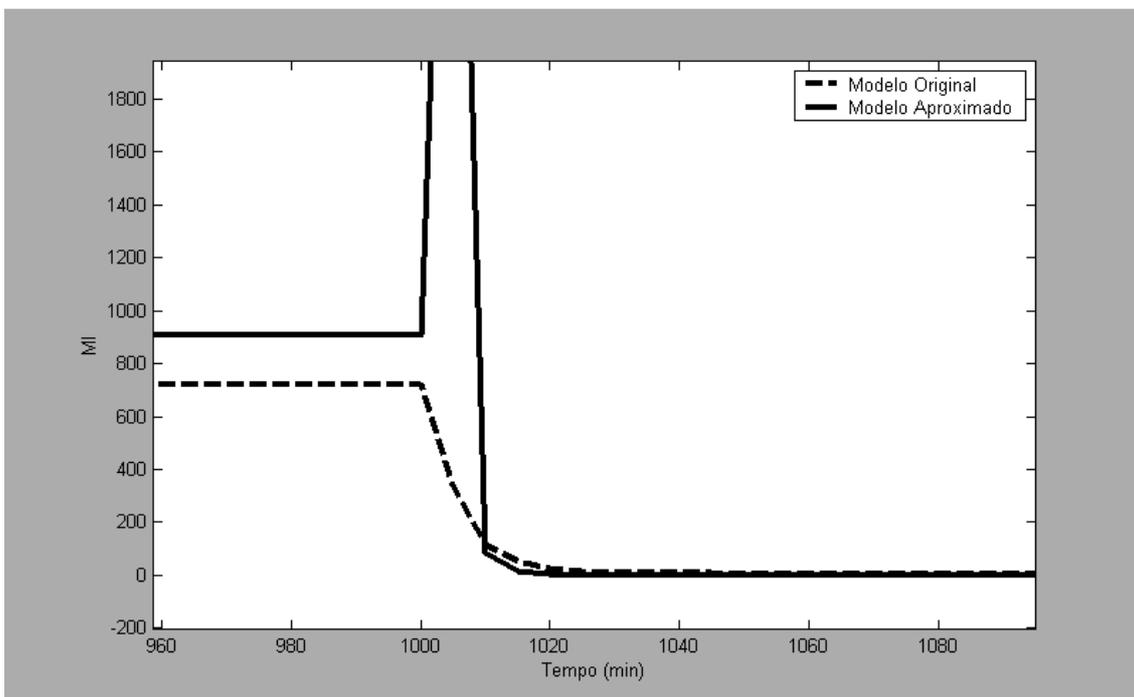
ver que os estados estacionários dos pontos da malha retangular foram corretamente obtidos e as divergências encontradas nas transições não presentes na identificação são conseqüências da interpolação.



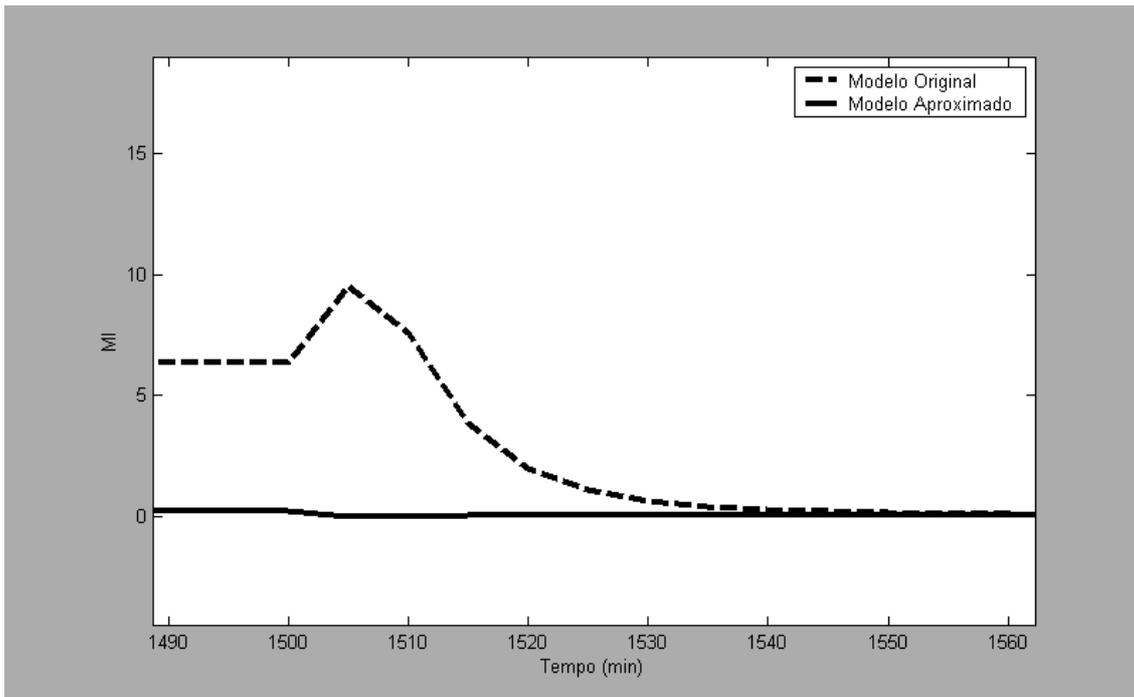
**Figure 5.19:** Comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem.



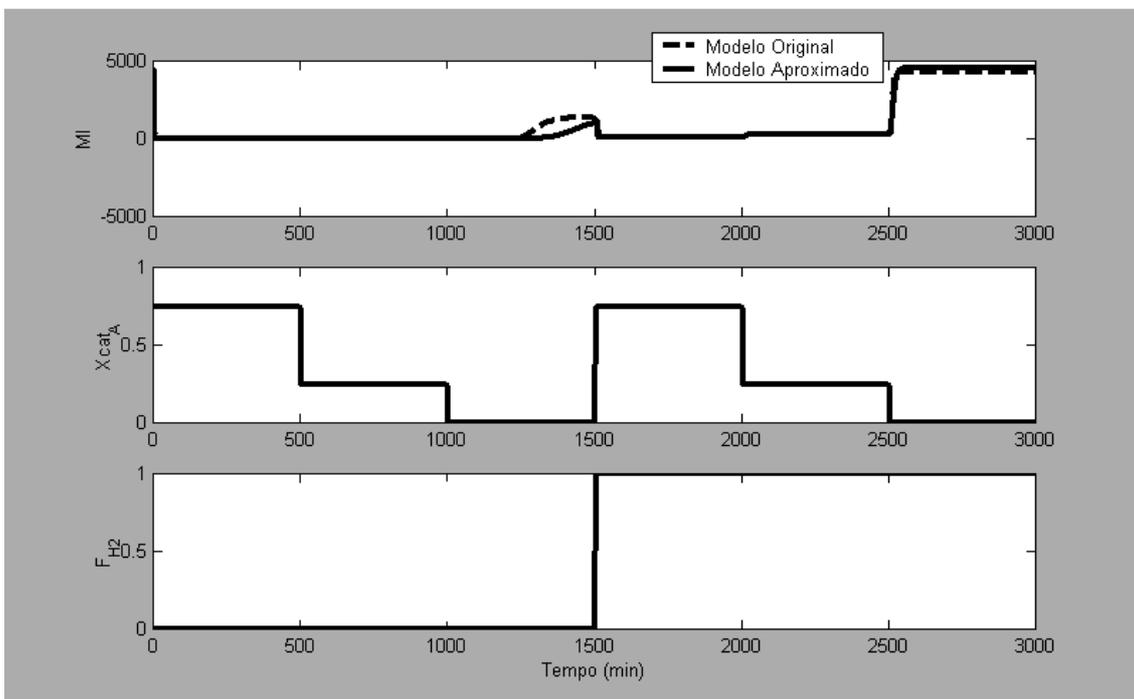
**Figure 5.20:** Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem.



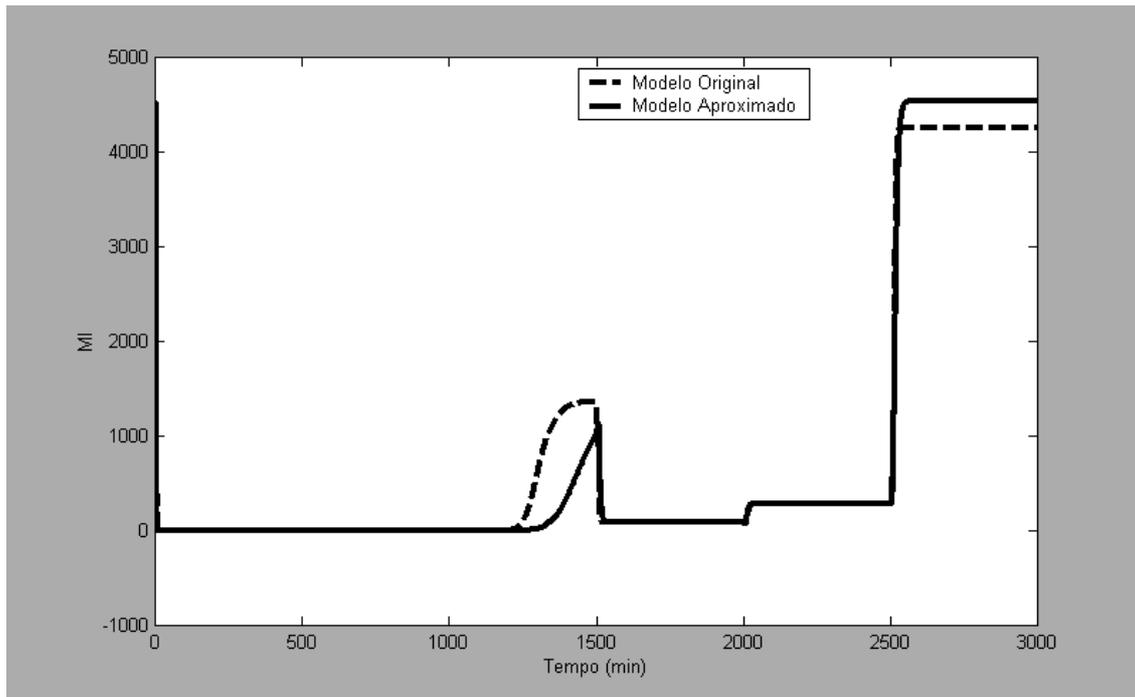
**Figure 5.21:** Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Terceira perturbação do conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem



**Figure 5.22:** Detalhes da comparação da resposta dos modelos original e aproximado para uma identificação baseada em momentos. Quarta perturbação do conjunto de perturbações diferentes das utilizadas na identificação do modelo aproximado. Modelos dos momentos de segunda ordem



**Figure 5.23:** Comparação das respostas do modelo original e aproximado para modelos baseados em momentos de segunda ordem frente a perturbações presentes no conjunto utilizado para a identificação.



**Figure 5.24:** Detalhe da comparação das respostas do modelo original e aproximado para modelos baseados em momentos de segunda ordem frente a perturbações presentes no conjunto utilizado para a identificação.

## 5.4 Otimização com Modelos Aproximados Simplificados

Utilizando-se a implementação do otimizador citada nas seções anteriores juntamente com o modelo simplificado de segunda ordem baseado em momentos, promoveu-se a otimização do processo do reator de polipropileno em simulação, ou seja, utilizando o modelo fenomenológico do reator como planta. Dentre os muitos testes realizados, aqui se destacam duas situações específicas que embasaram as conclusões deste trabalho. Na primeira situação, utilizou-se o otimizador apenas como controlador preditivo minimizando o erro entre o *setpoint* e a predição do modelo, além de um segundo termo de supressão de movimentos. Na segunda situação utilizou-se o algoritmo como otimizador propriamente dito avaliando-se um termo econômico que consistia basicamente em avaliar o custo do hidrogênio.

Na primeira situação os termos foram utilizados em uma estrutura multiobjetivo de atingimento de metas conforme mostrado no Capítulo 3. Tais termos são dados pelas equações:

$$(MI_{[k:k+P]} - MI_{set[k:k+P]})^2 - w_{MI} \cdot \gamma \leq 0 \quad (5.21)$$

$$\left(X_{catA[k:k+M]} - X_{catA[k-1:k+M-1]}\right)^2 - w_{X_{catA}} \cdot \gamma \leq 0 \quad (5.22)$$

$$\left(F_{H2[k:k+M]} - F_{H2[k-1:k+M-1]}\right)^2 - w_{FH2} \cdot \gamma \leq 0 \quad (5.23)$$

As Figuras 5.25 e 5.26 mostram a situação de mudanças de setpoint de MI próximo de zero para um MI igual a 10. Na situação da primeira figura a supressão de movimentos era menor que na segunda, levando a uma resposta mais oscilatória no primeiro caso. Note que nas duas situações o valor final atingido para cada variável manipulada é diferente. Isto ocorre porque devido à sobra de graus de liberdade, podem existir infinitas combinações das duas variáveis manipuladas que levam ao mesmo índice de fluidez. Note inclusive que na segunda situação a fração de catalisador A, que tem um ganho negativo com relação ao MI se analisado com uma vazão de hidrogênio fixa, é maior (0,82) que seu valor inicial (0,5). Isto ocorre porque a fração da vazão total de hidrogênio é maior na segunda situação (0,49) do que na primeira (0,28).

Esta sobra de graus de liberdade abre espaço para a otimização. Neste caso hipotético, será considerado como ótimo reduzir-se a vazão de hidrogênio ao mínimo fazendo com que o grade desejado de polímero seja função da combinação dos catalisadores. A vazão de hidrogênio neste caso específico deverá atuar apenas para transformar a resposta do índice de fluidez mais rápida, otimizando a transição, e retornando a um valor próximo de zero ao final da transição. Para tanto se adicionou o seguinte termo ao otimizador:

$$\left(F_{H2[k:k+M]}\right)^2 - w_{CFH2} \cdot \gamma \leq 0 \quad (5.24)$$

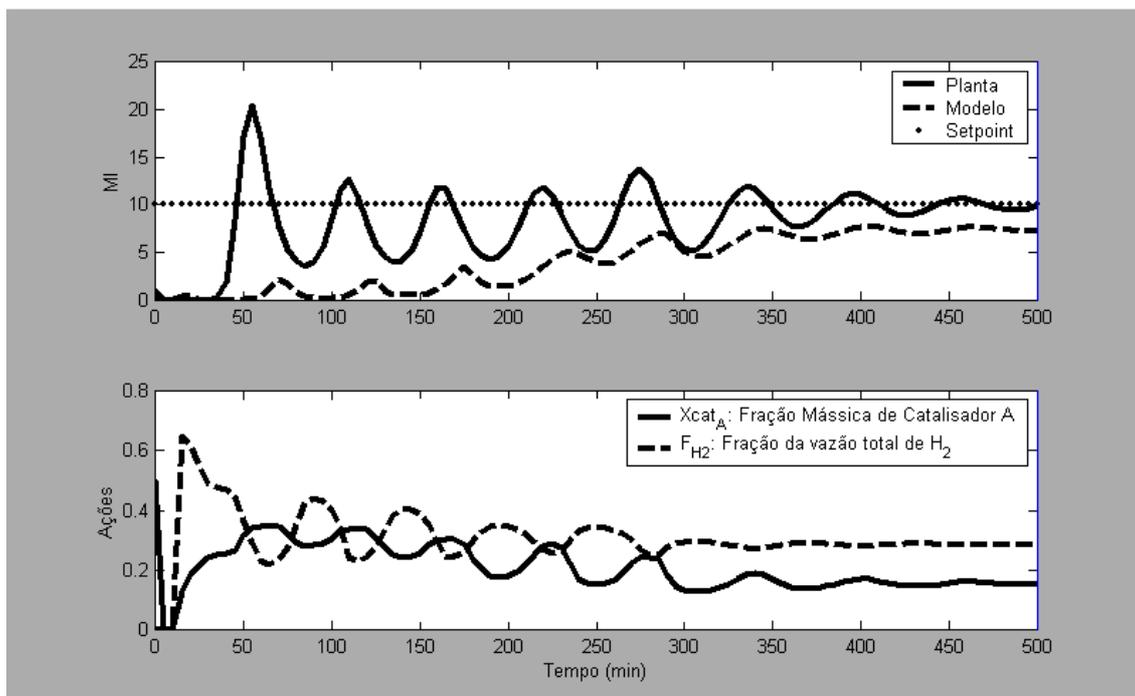
A resposta à mesma mudança de ponto de operação, anteriormente realizada sem otimização da vazão de hidrogênio, foi realizada novamente tentando-se otimizar tal vazão. O resultado de tal mudança pode ser visto nas Figuras 5.27 e 5.28. Cada uma das figuras, geradas com pesos diferentes do custo de hidrogênio, mostra que inicialmente o otimizador busca alcançar o ponto de operação desejado. No entanto, à medida que o erro entre predição e setpoint começam a tornar-se pouco significativos a parcela do custo do hidrogênio começa a trazer as duas variáveis manipuladas para zero com o intuito de operar com a menor vazão de hidrogênio possível e menor quantidade de catalisador A. No entanto, neste momento ocorre a instabilização do processo, cuja causa se deve a inversão do ganho do modelo

em uma região de baixa fração de catalisador A. Este comportamento é claramente visto ao observar-se as linhas do modelo aproximado e planta a partir do tempo 280 min na figura 5.27 e tempo 500 min na figura 5.28, onde observa-se que as duas linhas divergem do setpoint em direções opostas do setpoint.

Três parâmetros de ajustes estavam disponíveis nesta implementação. O primeiro destes parâmetros é o tempo de amostragem o qual consiste no intervalo entre duas amostragens realizadas pelo otimizador do processo. O segundo parâmetro de ajuste é o bloco de predição que consiste nos valores de tempo preditos pelo modelo que serão otimizados. O terceiro parâmetro é o Bloco de Ações de Controle que é os valores de tempo no futuro para os quais o otimizador projetará uma atuação no processo a fim de otimizá-lo. Os parâmetros de ajuste utilizados nesta implementação encontram-se na tabela 5.2

**Tabela 5.2:** Parâmetros de configuração do otimizador

Parâmetro	Valor (min)
Bloco de Ações de Controle	1 5 55 105 170
Bloco de Predição	1 5 55 105 170 275 445
Tempo de Amostragem	1



**Figura 5.25:** Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10.

$$w_{MI} = 1, w_{XcatA} = 0.1, w_{FH2} = 0.1$$

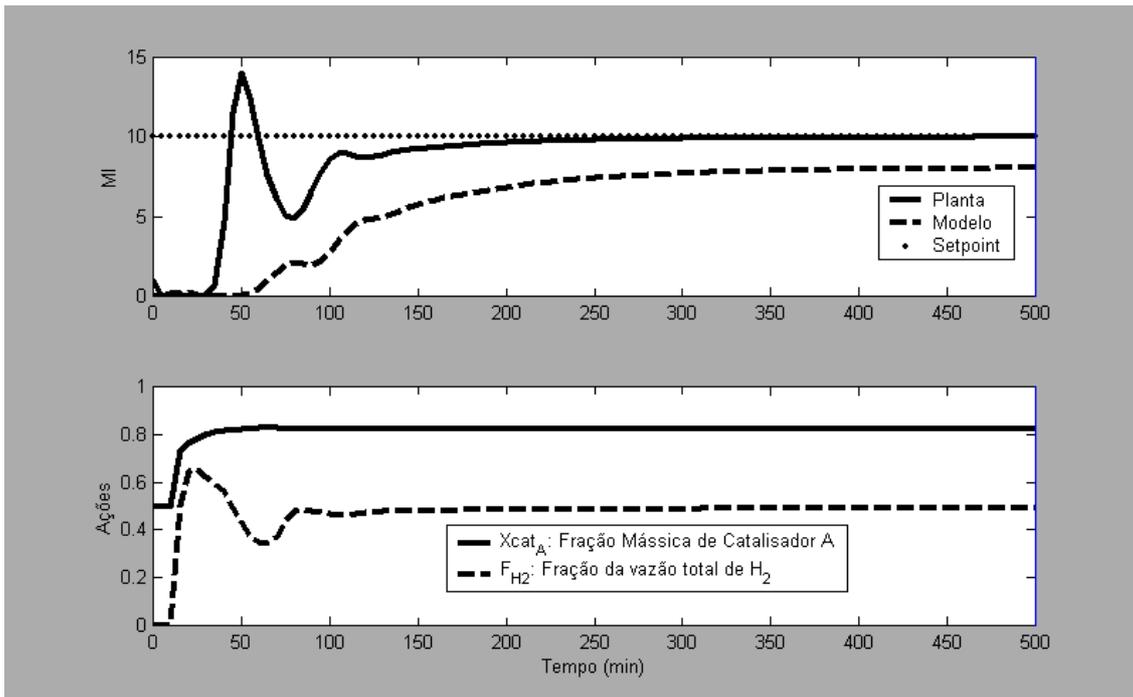


Figure 5.26: Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10.

$$w_{MI} = 3.3, w_{XcatA} = 0.1, w_{FH2} = 0.1$$

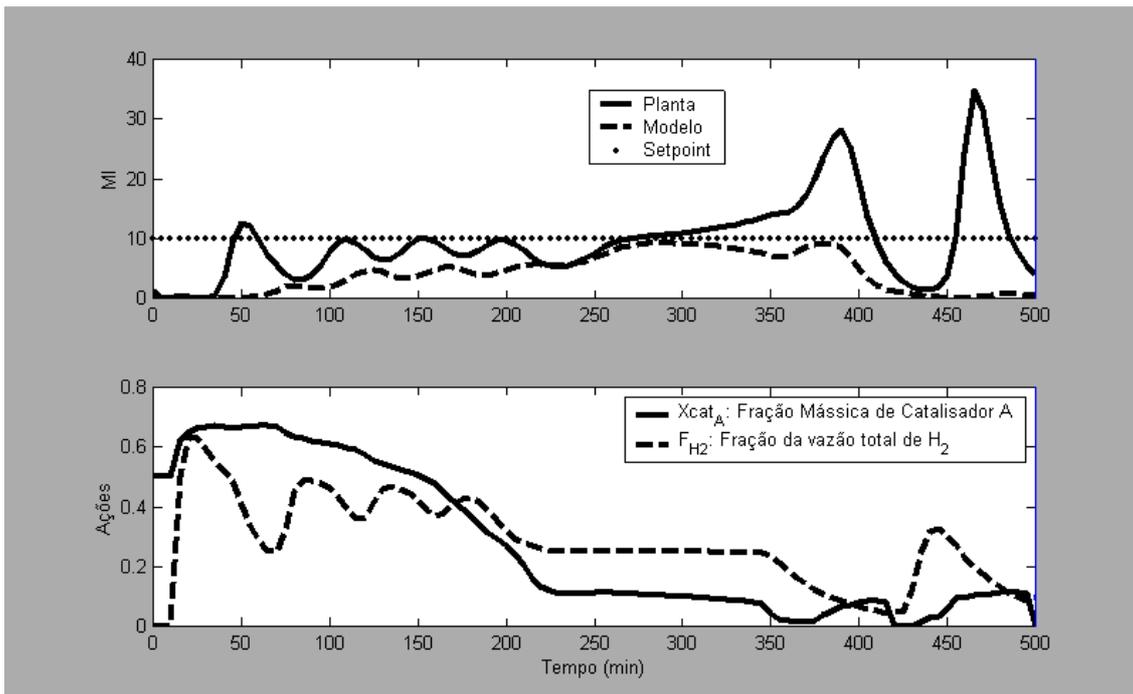
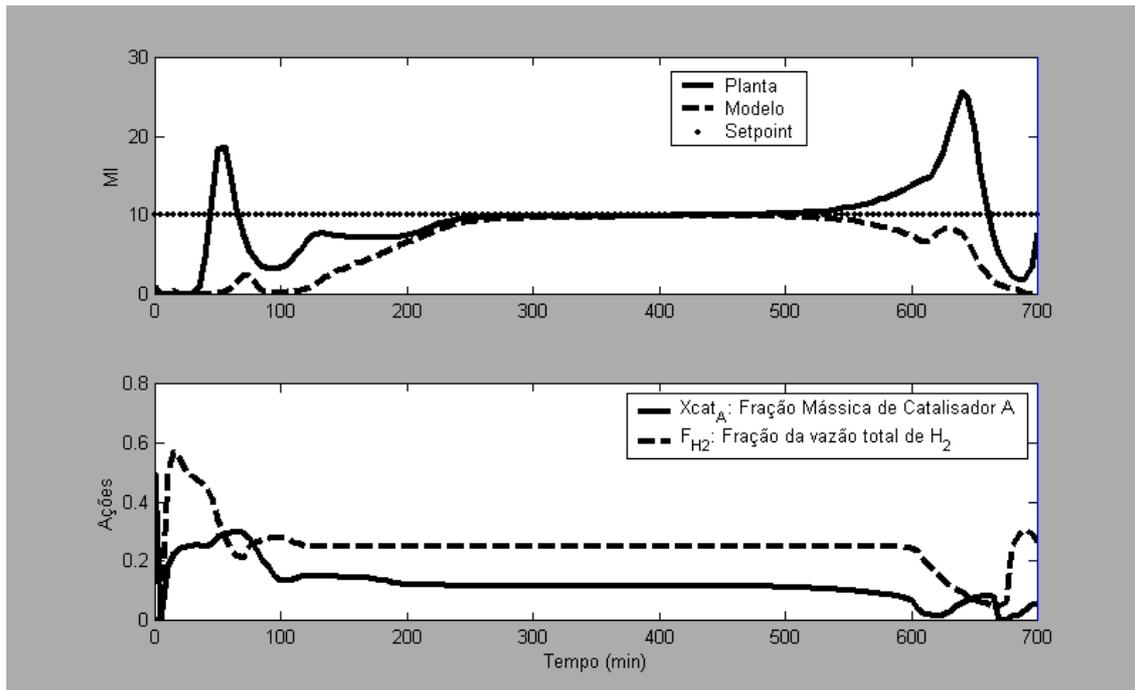


Figure 5.27: Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10.

$$w_{MI} = 3.3, w_{XcatA} = 0.1, w_{FH2} = 0.1, w_{CFH2} = 3.3$$



**Figure 5.28:** Resposta do otimizador frente a uma mudança de setpoint do MI de 0 para 10.

$$w_{MI} = 3.3, w_{X_{catA}} = 0.1, w_{F_{H_2}} = 0.1, w_{C_{FH_2}} = 10$$

## 5.5 Conclusões

Como considerações iniciais, é possível fazer-se uma análise do modelo aproximado obtido e seus impactos na aplicação do otimizador.

Observou-se que a variável de interesse na otimização (MI) apresenta um comportamento estacionário fortemente não linear em relação às variáveis manipuladas. A presença de tal não-linearidade degradou o desempenho da aplicação da conversão do modelo original em um aproximado com a metodologia até então desenvolvida. O comportamento do modelo aproximado para tal situação mostrou-se extremamente insatisfatório, além de não descrever o comportamento dinâmico adequado também não foi capaz de reproduzir adequadamente os pontos estacionários do modelo.

A alternativa encontrada para reproduzir com melhor qualidade o comportamento do índice de fluidez foi desenvolver modelos cujas saídas eram os momentos do polímero e, com estes, determinar variável de interesse (MI) usando as correlações não-lineares. Tal alternativa mostrou-se mais eficiente para representar os estados estacionários do modelo, mas não foi suficientemente eficaz para descrever as dinâmicas do modelo original. Melhorias na representação destas dinâmicas foram obtidas modificando-se a

ordem dos modelos de momentos, no entanto não se obteve uma descrição suficientemente boa.

Mesmo não tendo sido encontrada uma descrição aceitável, aplicou-se a otimização em tempo real do processo em simulação a fim de verificar se as não conformidades dos modelos seriam pontos limitantes para a sua aplicação e, após exaustivas simulações, verificou-se problemas em sua aplicação. Nas situações em que o otimizador foi utilizado como controlador preditivo obteve-se sucesso em sua aplicação, uma vez que a resposta na variável de interesse atingiu o setpoint e mostrou-se sensível aos ajustes do controlador da forma esperada. No entanto, a aplicação do otimizador com o intuito de promover uma transição otimizada levou o modelo e o processo a uma região onde a predição do modelo não era condizente com a predição do processo (modelo discrepante) o que provocou a instabilização do sistema. Esta constatação apontou para o fato de que a aplicação deste tipo de modelo simplificado na otimização de processos industriais ainda necessita de consideráveis desenvolvimentos para garantir a sua aplicação e indica que, nos moldes atuais, esta metodologia não é ainda aplicável.

A comparação do otimizador utilizando o modelo igual ao do processo não foi possível, devido ao tempo computacional necessário para realizar esta otimização. Enquanto na otimização com modelos simplificados o tempo requerido para se executar os cálculos de cada amostragem eram de 40s, o tempo necessário quando se utilizava o modelo igual ao do processo era de aproximadamente 7 horas. Isso impossibilitou a comparação tendo em vista que o tempo requerido para se fazer uma simulação completa, similar as realizadas neste capítulo seria de aproximadamente 3500h. Esta constatação mostra o ganho de tempo computacional que modelos simplificados podem trazer.

Os resultados obtidos com estes tipos específicos de modelos aplicados à otimização em tempo real não foram satisfatórios, mas indicam que desenvolvimentos futuros nesta área são extremamente necessários e podem trazer avanços consideráveis para a otimização em tempo real.

## 5.6 Referências Bibliográficas

Duraiski RG, Trierweiler JO, and Secchi aR. (2001). Nonlinear Model Predictive Control Using Successive Linearization Approach. ENPROMER - Encontro de Engenharia de Processo do Mercosul. Santa Fé - Argentina.

Salau. NPG, Secchi. AR, and Trierweiler JO. (2007). Five Formulations of Extended Kalman Filter: Which is the best for DRTO? In *European Symposium on Computer Aided Process Engineering – ESCAPE17*. Praga.



## Capítulo 6

### Conclusões

Neste capítulo, são apresentadas as conclusões obtidas no decorrer das diferentes frentes de trabalho desenvolvidas nesta tese. É possível dividir o trabalho em três etapas fundamentais, as quais foram detalhadas nos três últimos capítulos dados por:

- Aperfeiçoamento de um algoritmo de otimização:
- Proposição de uma estratégia de conversão de modelos rigorosos em versões simplificadas, de solução muito mais rápida que os originais:
- Aplicação do algoritmo de otimização e modelos convertidos na solução de um problema de otimização dinâmica em tempo real.

Embora os objetivos do trabalho não tenham sido atingidos em sua plenitude é possível destacar um conjunto significativo de melhorias e novas propostas de metodologias úteis para a aplicação na engenharia. Na seção seguinte serão detalhados tanto os sucessos como os insucessos obtidos neste trabalho.

#### 6.1 Considerações Gerais

Quanto ao algoritmo desenvolvido, a proposta original era a obtenção de um algoritmo que permitisse um aumento na velocidade de resolução de problemas de otimização e tornasse possível a aplicação de um otimizador em tempo real com

uma frequência de execução alta o suficiente para permitir a sua aplicação na forma mono camada juntamente com o controle preditivo.

É possível destacar como pontos fortes do desenvolvimento a obtenção de uma metodologia bastante eficiente para a escolha de variáveis dependentes e independentes de cada passo das iterações do algoritmo além da obtenção de uma direção de busca otimizada, a qual procura seguir na direção de maior redução do valor função. Os cálculos do algoritmo GA-GRG que, à primeira vista, parecem ser mais complexos e tomar um longo tempo para sua execução, apresentaram uma efetividade melhor que a da solução do algoritmo SQP, indicando que o tempo de solução do algoritmo em muitos casos é menor que o do algoritmo SQP. Por outro lado, o algoritmo SQP estudado mostrou uma maior robustez, melhor qualidade, e melhor eficiência que o algoritmo GA-GRG, o que mostra uma capacidade de resolução pior do algoritmo GA-GRG, uma menor precisão da solução, e exigindo mais avaliações da função objetivo.

A falta de precisão numérica na solução não é um problema grave para a otimização de processos industriais uma vez que, fazendo-se um escalonamento adequado, transforma-se as variáveis em intervalos adequados onde incertezas na quarta casa decimal não são tão importantes. A falta de robustez encontrada no algoritmo GA-GRG deve-se basicamente a avaliações de restrições bilineares as quais modificam a sua sensibilidade em relação à importância das variáveis escolhidas como independentes na direção de busca, impossibilitando a solução por meio deste algoritmo nos moldes atuais. Já o excesso do número de avaliações da função objetivo é um problema se a otimização acontecer avaliando os modelos fenomenológicos, os quais tomam um considerável tempo computacional para a sua solução.

Não foi obtida uma solução para o problema da robustez do algoritmo GA-GRG nesta tese. Já o número de avaliações excessivo da função objetivo torna-se pouco importante quando se trabalha com modelos simplificados tal como o que está sendo proposto nesta tese.

Embora a velocidade de execução do algoritmo tenha apresentado uma melhora em relação ao SQP apresentado, o ganho de tempo obtido não foi suficiente para tornar a solução do otimizador tão rápida quanto se queira, uma vez que o objetivo final é a utilização deste na forma monocamada. Isto ocorre porque se percebeu que o grande responsável pela demora na solução de problemas de

otimização dinâmica e a avaliação dos modelos de grande dimensão baseados em equações diferenciais. Esta constatação motivou o desenvolvimento dos estudos mostrados no Capítulo 4 referentes à geração dos modelos simplificados.

A idéia por trás dos modelos simplificados era a obtenção de modelos que descrevessem o comportamento não linear do processo com uma boa aproximação da realidade, capazes de serem resolvidos em frações de segundos. Embora se tenha obtido modelos extremamente rápidos, os mesmos apresentaram problemas quanto à descrição do modelo original.

A grande deficiência do método de conversão de modelos fenomenológicos em simplificados decorre da consideração inicial de que a dinâmica da transição seja dependente única e exclusivamente do estado estacionário final. No Capítulo 4 é possível observar-se que a dinâmica na realidade depende também do ponto final da perturbação e da direção de onde este se inicia. Não foi possível obter-se um novo método que levasse em conta a direção da perturbação porque a concatenação dos estados de diferentes modelos lineares obtidos em dois pontos diferentes da transição era dependente desta prerrogativa.

Observa-se que as conversões do modelo de seis tanques esféricos mostrados no Capítulo 4 apresentaram estados estacionários condizentes com os do modelo original. No entanto, as dinâmicas obtidas apresentaram consideráveis divergências em relação às encontradas no modelo original, provavelmente, em função da não consideração da direção de perturbação.

Assim sendo, surge aqui uma possibilidade de desenvolvimento futuro para esta metodologia – a obtenção de uma forma de conversão capaz de levar em consideração, não apenas o estado estacionário final, mas também a direção da perturbação que leva o modelo a este ponto. Caso seja possível encontrar a relação correta da direção de perturbação com a dinâmica do modelo é possível que se obtenha um modelo dinâmico com alta reprodutibilidade e extremamente rápido para utilização em aplicativos online.

Outro problema relativo à conversão de modelo utilizando a metodologia desenvolvida aqui é a falta de capacidade de se trabalhar com modelos integrativos, instáveis, ou com pólos complexos, pois tais casos não estão previstos no método. Além disso, a obtenção das constantes de tempo é dependente do tempo de amostragem escolhido para a coleta de dados do sistema. No entanto, pode-se destacar como ponto extremamente positivo do método a obtenção de uma forma de

converter modelos de resposta ao degrau em modelos no espaço de estado e função de transferência, sem a necessidade de qualquer tipo de otimização.

Durante a aplicação de otimização da transição de pontos de operação mostrada no capítulo anterior observou-se outro problema, bastante grave e restritivo para a aplicação da metodologia de conversão de modelos, mas que pode ser contornada no caso específico estudado – a não reprodutibilidade dos estacionários originais no modelo interpolado para variáveis de saída muito não-lineares. No caso específico, o problema foi contornado pela utilização de variáveis base da representação da variável de interesse, as quais apresentavam um comportamento bem mais linear do que a variável desejada, fazendo-se a posterior conversão estática das variáveis base na variável de interesse. Neste caso a falta de reprodutibilidade do estado estacionário ocorreu devido ao fato de se utilizar uma interpolação cúbica dos estados estacionários, o que limita consideravelmente as possibilidades de representação não linear. Embora neste caso o problema tenha sido contornado, nada garante que em outros casos de aplicação serão encontradas variáveis base suficientemente linear para permitir a aplicação da mesma solução encontrada, o que torna a aplicação deste método limitada também neste sentido.

Novamente na aplicação ao exemplo de otimização mostrou-se a necessidade da escolha correta da ordem do modelo aproximado a fim de encontrar uma representação que não corte as dinâmicas importantes do modelo original e nem adicione comportamentos inexistentes. Ainda que se escolha a dinâmica mais adequada, isso não é garantia de que a representação encontrada seja boa justamente pelo fato de não se capturar todos os efeitos da direcionalidade do sistema.

Por fim, observa-se que para algumas regiões de operação do otimizador obteve-se sucesso na otimização da transição e o comportamento encontrado era condizente com o esperado. No entanto, encontraram-se regiões em que o ganho descrito pelo modelo original não condizia com o ganho do modelo aproximado, o que provocou a instabilização do sistema. Este comportamento incerto do modelo aproximado leva a conclusão de que, nos moldes atuais, ainda não é possível aplicar-se esta metodologia de otimização em tempo real por dois principais motivos:

- A falta de uma descrição apropriada que capture os efeitos da direcionalidade do modelo;

- A falta de uma metodologia de interpolação capaz de gerar modelos com ganhos condizentes com o do modelo real.

Os problemas encontrados anteriormente podem servir de impulsionadores para trabalhos futuros que se proponham a resolvê-los, pois embora a aplicação de tal metodologia não tenha dado bons resultados, a idéia de se obter modelos dinâmicos de rápida avaliação pode ser a solução de muitos problemas de aplicações online dependentes de modelo e que hoje esbarram na limitação de capacidade de processamento computacional.



## **Anexo I**

**Relatório: “Modelagem de Polimerização de Polipropileno via Catalisador Ziegler-Natta em reator CSTR slurry”**



# Modelagem de Polimerização de Polipropileno via Catalisador Ziegler-Natta em reator CSTR *slurry*<sup>1</sup>

## 1. Descrição Cinética

A descrição cinética considerada a seguir corresponde àquela encontrada nos trabalhos de Secchi e Coni Jr. (referências [1] e [9]). Envolve uma distribuição de sítios ativos e a diferenciação entre as cinéticas de geração de polímero atático e isotático, bem como a desativação do catalisador e a reação de transferência para o hidrogênio. Assume-se neste trabalho que as constantes cinéticas dependem da temperatura na forma da Lei de Arrhenius:

$$K_{i,j} = A_{i,j} \exp\left(-\frac{E_{i,j}}{RT}\right), \text{ onde o subscrito 'i' se refere à reação e o subscrito 'j'}$$

ao tipo de sítio ativo.

---

<sup>1</sup> Modelo utilizado desenvolvido por *Pedro Rafael Bolognese Fernandes* no âmbito da disciplina Modelagem e Simulação de Reatores de Polimerização oferecida pelo PPGEQ/UFRGS.

A dependência com a pressão foi desprezada. Por simplicidade, também, admitiu-se que o fator pré-exponencial é idêntico para todos os sítios em uma dada reação:  $A_{i,j} = A_i$ .

### 1.1. Distribuição de Sítios

A energia de ativação da reação ‘ $i$ ’ no sítio ‘ $j$ ’ é considerada como uma distribuição discreta que é função da variável reduzida  $z_j$  e tem média e desvio-padrão conhecidos:

$$E_{i,j} = \bar{E}_i + \sigma_i z_j$$

A função geratriz da variável reduzida é considerada gaussiana:

$$z_j = \frac{\phi(j+1)}{ns+1} - 3$$

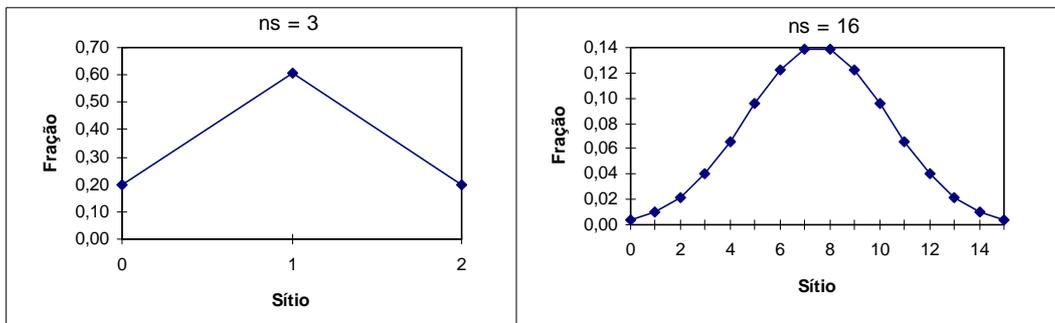
onde  $ns$  é o número de tipos de sítios ativos diferentes. Os índices ‘ $j$ ’ variam, portanto, de 0 até  $ns-1$ . Ao contrário dos dois trabalhos mencionados acima, a fração de cada tipo de sítio não é dada pela distribuição acumulada (probabilidade) da função gaussiana (que é contínua por origem), mas é substituída por uma função discreta que se ajusta a uma densidade de probabilidade normal em função do número de sítios:

$$f_n = K \exp\left(-\frac{n^2}{2}\right), \text{ para } K = \left(\sum_{n=0}^r \exp\left(-\frac{n^2}{2}\right)\right)^{-1} \quad (n \text{ é a variável reduzida}).$$

Esta alteração não provoca diferenças significativas, mas apresenta maior coerência. A comparação a seguir ilustra o resultado:

Fração de sítios para $ns = 3$		
Sítio	Este trabalho	Coni Jr.
1	0.197	0.188
2	0.606	0.624
3	0.197	0.188

Os gráficos a seguir mostram o resultado da distribuição discreta para dois casos. Notar que a mesma é normalizada, ou seja, o somatório das frações (que é a probabilidade acumulada total) é unitário.



**Gráfico 1.** Distribuição normal discreta proposta neste trabalho.

## 1.2. Reações Consideradas

A seguinte notação é empregada a seguir para representar as espécies envolvidas; o subscrito ‘*j*’ indica o tipo de sítio, o subscrito ‘*r*’ indica o grau de polimerização e o superescrito ‘*a*’ indica ataticidade. A ausência deste significa isotaticidade.

$M$  representa o monômero;

$C_j, C_j^a$  representa um sítio não ativado;

$S_j, S_j^a$  representa um sítio ativo isotático/atático do tipo ‘*j*’;

$L_{r,j}, L_{r,j}^a$  representa um cadeia de polímero vivo;

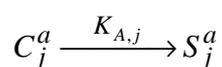
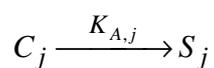
$D_{r,j}, D_{r,j}^a$  representa um cadeia de polímero morto;

$X_j^a, X_j$  representa um sítio desativado.

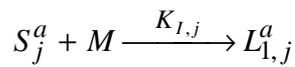
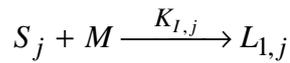
$N$  indica o doador de elétrons (‘*donor*’).

As reações consideradas são então:

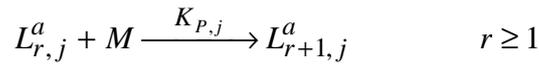
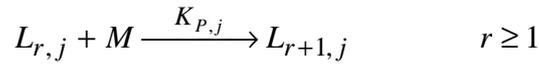
### I. Ativação Espontânea de Sítios:



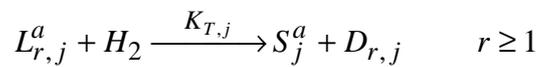
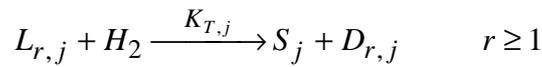
II. Iniciação:



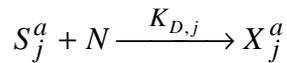
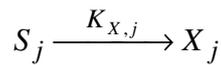
III. Propagação:



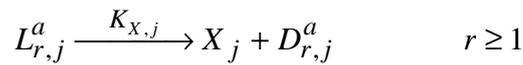
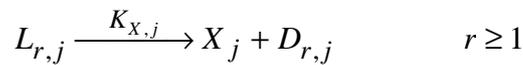
IV. Transferência para o hidrogênio:



V. Desativação de Sítios:



VI. Terminação Espontânea:



As reações indicadas ocorrem na fase polímero, exceto as de ativação de sítio. No entanto, para efeito de modelagem, as espécies serão consideradas homoganeamente distribuídas no meio reacional (líquido), portanto, desprezar-se-ão os questões relativas à difusão.

## 2. Expressões de Taxa

As taxas de reação serão consideradas em função do produto entre as concentrações molares de cada participante da reação. Assim, todas as constantes de taxa serão dadas em unidades de [volume/mol.tempo], exceto  $K_{A,j}$  e  $K_{X,j}$ , dadas em unidades de [1/tempo]. A taxa volumétrica de consumo de monômero será, portanto:

$$(-r_M) = K_{I,0}M(S_0 + S_0^a) + K_{I,1}M(S_1 + S_1^a) + \dots + K_{P,0}M(L_{1,0} + L_{1,0}^a) + K_{P,1}M(L_{1,1} + L_{1,1}^a) + \dots \\ + K_{P,0}M(L_{2,0} + L_{2,0}^a) + K_{P,1}M(L_{2,1} + L_{2,1}^a) + \dots + K_{P,0}M(L_{3,0} + L_{3,0}^a) + K_{P,1}M(L_{3,1} + L_{3,1}^a) + \dots$$

cujas unidades são:  $\left[ \frac{\text{mol}}{\text{volume} \times \text{tempo}} \right] = \left[ \frac{\text{volume}}{\text{mol} \times \text{tempo}} \right] \times \left[ \frac{\text{mol}}{\text{volume}} \right] \times \left[ \frac{\text{mol}}{\text{volume}} \right]$

A expressão da taxa acima, empregando-se o conceito de momento, pode ser escrita como:

$$(-r_M) = M \sum_{j=0}^{ns-1} K_{I,j}(S_j + S_j^a) + M \sum_{j=0}^{ns-1} \sum_{r=1}^{\infty} K_{P,j}(L_{r,j} + L_{r,j}^a) = M \sum_{j=0}^{ns-1} K_{I,j}(S_j + S_j^a) \\ + M \sum_{j=0}^{ns-1} K_{P,j}(\mu_{0,j}^L + \mu_{0,j}^{L,a})$$

As expressões de taxa são usadas nas equações de balanço da seguinte maneira, por exemplo, para o monômero:

$$\frac{d(MV)}{dt} = F_{M,in} - F_M + (-r_M)V \quad \text{onde } F_M \text{ representa a vazão molar de}$$

saída de monômero.

### 3. Expressões do Balanço Material

As variáveis empregadas nos balanços serão notadas como na listagem a seguir.

As concentrações são todas em unidade de [mol/L]:

$M$ : concentração de monômero;

$E$ : concentração de solvente;

$S_j$ : concentração de sítio ativo do tipo 'j';

$L_r$ : concentração de polímero vivo com tamanho de cadeia 'r';

$D_r$ : concentração de polímero morto com tamanho de cadeia 'r';

$$\mu_{k,j}^L: \text{momento de ordem k do polímero vivo: } \mu_{k,j}^L = \sum_{r=1}^{\infty} r^k L_{r,j}$$

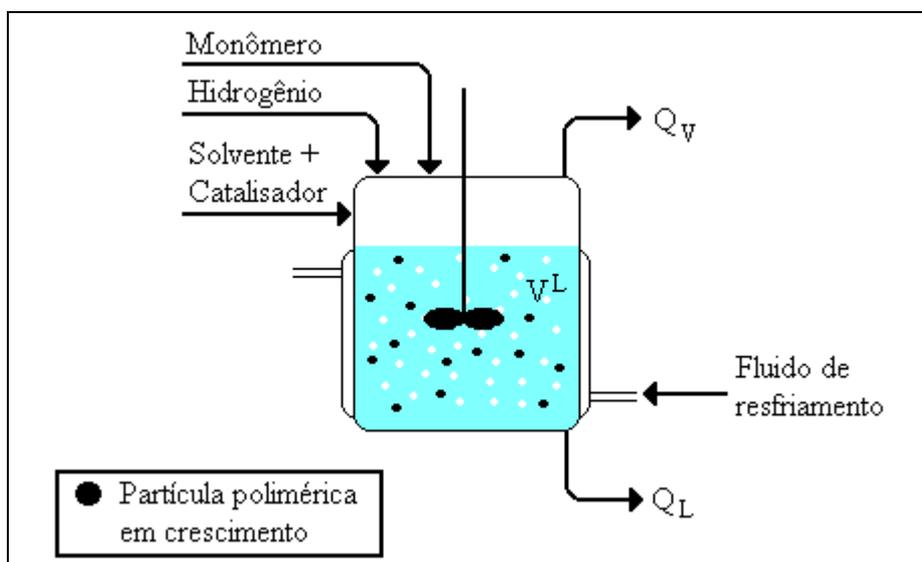
$$\mu_k^D: \text{momento de ordem k do polímero morto: } \mu_k^D = \sum_{r=1}^{\infty} r^k D_r$$

$H_2$ : concentração de hidrogênio;

$F_M, F_{H_2}, F_{S_j}$ : vazões molares de  $M, H_2$  e  $S_j$ , em [mol/min].

$V$ : volume (em L).

Também será empregada a seguinte convenção: superescrito 'L' indica fase líquida e superescrito 'V' indica fase vapor; portanto, o superescrito aplicado à concentração indica a concentração de cada fase, e o mesmo vale para o volume. A ausência denota concentração total. Para o volume, o superescrito 'R' indica o volume total do reator. Ainda, o subscrito 'in' aplicado às vazões molares indica vazão de entrada. A figura a seguir representa a configuração considerada.



**Figura 2.** Reator *slurry* CSTR de polimerização de propileno modelado.

O balanço material para o reator CSTR em operação é dado a seguir;  $q_L$  é a vazão volumétrica de retirada de líquido do reator, e  $q_v$  a de retirada de vapor:

$$\frac{d(MV^R)}{dt} = F_{M,in} - M^L \cdot q_L - M^V \cdot q_v - M^L \sum_{j=0}^{ns-1} \left[ K_{I,j} (S_j + S_j^a) + K_{P,j} (\mu_{0,j}^L + \mu_{0,j}^{L,a}) \right] V^L$$

$$\frac{d(H_2 V^R)}{dt} = F_{H_2,in} - H_2^L \cdot q_L - H_2^V \cdot q_v - H_2^L \sum_{j=0}^{ns-1} \left[ K_{T,j} (\mu_{0,j}^L + \mu_{0,j}^{L,a}) \right] V^L$$

$$\frac{d(S_j V^L)}{dt} = F_{S_j,in} - S_j \cdot q_L - \left[ (K_{I,j} M + K_{X,j}) S_j - K_{T,j} H_2 \mu_{0,j}^L \right] V^L$$

$$\frac{d(S_j^a V^L)}{dt} = F_{S_j^a,in} - S_j^a \cdot q_L - \left[ (K_{I,j} M + K_{D,j} N) S_j^a - K_{T,j} H_2 \mu_{0,j}^{L,a} \right] V^L$$

$$\frac{d(L_{1,j} V^L)}{dt} = \left[ M(K_{I,j} S_j - K_{P,j} L_{1,j}) - (K_{T,j} H_2 + K_{X,j}) L_{1,j} \right] V^L - L_{1,j} \cdot q_L$$

$$\frac{d(D_1 V^L)}{dt} = -D_1 \cdot q_L + \sum_{j=0}^{ns-1} \left[ (K_{T,j} H_2 + K_{X,j}) L_{1,j} \right] V^L$$

$$\frac{d(\mu_{0,j}^L V^L)}{dt} = \left[ M K_{I,j} S_j - (K_{T,j} H_2 + K_{X,j}) \mu_{0,j}^L \right] V^L - \mu_{0,j}^L \cdot q_L$$

$$\frac{d(\mu_{0,j}^{L,a} V^L)}{dt} = \left[ M K_{I,j} S_j^a - (K_{T,j} H_2 + K_{X,j}) \mu_{0,j}^{L,a} \right] V^L - \mu_{0,j}^{L,a} \cdot q_L$$

$$\frac{d(\mu_{k,j}^L V^L)}{dt} = \left[ M \left( K_{I,j} S_j + K_{P,j} \sum_{i=1}^k C_{k,i} \mu_{k-i,j}^L \right) - (K_{T,j} H_2 + K_{X,j}) \mu_{k,j}^L \right] V^L - \mu_{k,j}^L \cdot q_L$$

onde  $ng \geq k \geq 1$

$$\frac{d(\mu_k^D V^L)}{dt} = -\mu_k^D \cdot q_L + \sum_{j=0}^{ns-1} (K_{T,j} H_2 + K_{X,j}) \mu_{k,j}^L V^L \quad \text{onde } ng \geq k \geq 0$$

$$\frac{d(\mu_0^{D,a} V^R)}{dt} = -\mu_0^{D,a} \cdot q_L + \sum_{j=0}^{ns-1} (K_{T,j} H_2 + K_{X,j}) \mu_{0,j}^{L,a} V^L$$

$$\frac{d(NV^L)}{dt} = F_{N,in} - N \cdot q_L - N \sum_{j=0}^{ns-1} \left[ K_{D,j} S_j^a \right] V^L$$

As concentrações de sítios, ‘donor’ e aquelas referentes ao polímero, que ocorrem somente na fase líquida, foram referidas ao volume desta, apesar de poder-se ter usado o volume total do reator.

As condições iniciais necessárias para a resolução do sistema de equações diferenciais são, por exemplo,  $M(0) = M_0$ ,  $H_2(0) = H_{2,0}$ , sendo estas duas referentes à fase líquida,  $S_j(0) = S_0$ , onde  $S_0$  é a concentração de sítios ativos do tipo ‘j’. As demais condições iniciais para a simulação dinâmica são análogas.

O sistema de equações diferenciais pode ser escrito da seguinte maneira compacta, através de vetores facilmente identificáveis acima:

$$\frac{d(\tilde{C}V)}{dt} = \tilde{F}_{in} + (\tilde{r}_g)V - \tilde{F}$$

Escrevendo para o componente ‘  $i$  ’, teremos, então:

$$\frac{d(C_i V)}{dt} = F_{i,in} - F_i + \sum_{j=1}^m (r_{i,j})V$$

onde o somatório indica a taxa de formação do componente na reação ‘  $j$  ’, (obviamente, só daquelas em que participa). Uma vez que as taxas numa mesma reação podem ser relacionadas pelos coeficientes estequiométricos  $\nu_i$ , positivos para produtos e negativos para reagentes, escrevemos agora para um conjunto de ‘  $r$  ’ reações:

$$\frac{d(C_i V)}{dt} = F_{i,in} - F_i + \sum_{j=1}^r \nu_{i,j} (-r_j) V \quad (1)$$

onde  $\nu_{i,j}$  é o coeficiente estequiométrico do componente ‘  $i$  ’ na reação ‘  $j$  ’, igual a  $\pm 1$  se participar da mesma, e 0 caso contrário. A taxa de reação foi tomada com relação a um dos reagentes da reação (portanto, temos ‘  $r$  ’ taxas).

As considerações que permitem escrever-se os balanços transientes acima são de que as concentrações de saída são idênticas às encontradas no interior do reator. Apesar da descrição mais apropriada considerar a existência de fases dispersas no reator, e, portanto, de fenômenos de transferência de massa e energia nas interfaces, o modelamento aqui apresentado considera concentração uniforme no reator (CSTR).

Poder-se-ia ter empregado uma vazão molar de saída,  $F$  [mol/tempo]. Nesse caso, a taxa molar retirada de cada espécie seria proporcional à fração molar desta dentro do reator. De qualquer modo, as variáveis citadas estão relacionadas através da seguinte expressão:

$$F = q_L \times \sum_i C_i = q_L \times C_{total}, \text{ e, para o componente, } F_i = q_L \times C_i$$

Uma vez que o volume reacional não é, em princípio, constante, ainda falta o balanço global de massa a fim de determinar o mesmo:

$$\frac{d(CV^L)}{dt} = F_{M,in} + F_{H_2,in} + F_{E,in} - F \quad (2)$$

onde  $C$  é a concentração total de espécies presentes na fase líquida:  $C_{total} = \sum C_i = M + H_2 + \mu_{0,j}^L + \mu_0^D + \dots$

Finalmente, restam os termos de saída. Por simplicidade, considera-se aqui que um controle do tipo apenas proporcional pois as variáveis são acúmulos:

$$q_L = K_{P,1}(V^L - V_{set}^L)$$

$$q_V = K_{P,2}(P - P_{set})$$

### 3.1. Considerando as Fases Líquida e Vapor

Para o estudo da solubilidade de separação entre fases, separou-se os balanços de massa para o monômero, hidrogênio e solvente nas fases líquida e gasosa. A inclusão da equação referente a este último foi necessária a fim de poder-se calcular o volume da mistura reacional, que, em princípio, não é constante.

#### 3.1.1. Monômero:

Fase Líquida:

$$\frac{d(MV^L)}{dt} = F_{M,in}^L - M^L \cdot q_L - M^L \sum_{j=0}^{ns-1} \left[ K_{I,j}(S_j + S_j^a) + K_{P,j}(\mu_{0,j}^L + \mu_{0,j}^{L,a}) \right] V^L$$

Fase Vapor:

$$\frac{d(MV^V)}{dt} = F_{M,in}^V - M^V \cdot q_V$$

#### 3.1.2. Hidrogênio:

Fase Líquida:

$$\frac{d(H_2 V^L)}{dt} = F_{H_2,in}^L - H_2^L \cdot q_L - H_2^L \sum_{j=0}^{ns-1} \left[ K_{T,j}(\mu_{0,j}^L + \mu_{0,j}^{L,a}) \right] V^L$$

Fase Vapor:

$$\frac{d(H_2 V^V)}{dt} = F_{H_2,in}^V - H_2^V \cdot q_V$$

#### 3.1.3. Solvente:

Fase Líquida:

$$\frac{d(EV^L)}{dt} = F_{E,in}^L - E^L \cdot q_L$$

Fase Vapor:

$$\frac{d(EV^V)}{dt} = F_{E,in}^V - E^V \cdot q_V$$

As vazões de entrada em cada fase serão relacionadas adiante ao equilíbrio termodinâmico. Por questões de eficiência numérica, ao invés do produto entre concentração e volume de cada fase foi implementado o número de moles de cada espécie, em cada fase.

### 3.2. Equacionamento do Equilíbrio Termodinâmico

As equações que descrevem o equilíbrio estacionário de fases (*'flash'*) são as seguintes:

$$\text{Balanço de massa: } z_i - \beta y_i - (1 - \beta)x_i = 0$$

$$\text{Balanço de Energia: } h_q - \beta H - (1 - \beta)h = 0$$

$$\text{Somatório das Frações Molares: } \sum_{i=1}^n x_i - \sum_{i=1}^n y_i = 0$$

$$\text{Equilíbrio de Fases: } y_i - K_i x_i = 0$$

Onde  $z$ ,  $y$  e  $x$  são as composições molares da alimentação, da fase gás e da fase líquida, respectivamente,  $\beta$ , a fração vaporizada,  $h$  a entalpia da fases líquida e  $H$  a da fase gás. O grandeza  $K$  é conhecida como razão de equilíbrio da vaporização e é relacionada à pressão, temperatura e composição do sistema. Desse modo, podemos escrever:

$$F_{i,in}^L = (1 - \beta)x_i F \quad \text{e também} \quad F_{i,in}^V = \beta y_i F \quad . F \text{ é a alimentação molar total.}$$

### 4. Desenvolvimento das Expressões do Balanço de Energia

A aplicação da segunda lei da Termodinâmica ao sistema em fluxo considerando desprezáveis as parcelas de energia decorrentes de potenciais (gravitacional, elétrico e magnético) traz que:

$$\frac{d(E_{tot})}{dt} \approx \frac{dU}{dt} = \dot{Q} - \dot{W} + \sum F_{i,in} U_{i,in} - \sum F_i U_i = \dot{Q} - \dot{W}_S + \sum F_{i,in} H_{i,in} - \sum F_i H_i$$

Onde o subscrito 'in' denota as grandezas avaliadas na entrada do sistema,  $\dot{W}_S$  é o trabalho de eixo,  $U$  representa a energia interna e  $H$  a entalpia; a relação entre ambas é dada pelo trabalho reversível de compressão do elemento de volume:  $H = U + Pv$ , onde  $v$  representa o volume molar da mistura.  $\dot{Q}$  é a taxa de calor fornecida pelo ambiente ao reator, por exemplo, através de uma camisa de resfriamento. A energia (interna) total do sistema é soma das energias molares dos componentes, representados pelo número de moles  $n_i$ ; substituindo-se agora a derivada temporal da energia interna pela entalpia:

$$\frac{d\left(\sum_i n_i U_i\right)}{dt} = \frac{d\left(\sum_i n_i (H_i - Pv_i)\right)}{dt} = \dot{Q} - \dot{W}_S + \sum F_{i,in} H_{i,in} - \sum F_i H_i$$

Expandindo a derivada acima e desprezando a variação da pressão no tempo:

$$\sum_i n_i \frac{dH_i}{dt} + \sum_i H_i \frac{dn_i}{dt} = \sum_i (C_i V) c_{p,i} \frac{dT}{dt} + \sum_i H_i \frac{d(C_i V)}{dt} = \dot{Q} - \dot{W}_S + \sum F_{i,in} H_{i,in} - \sum F_i H_i$$

O termo da esquerda, uma vez que é avaliado dentro do volume de controle, se identifica com a saída do sistema. Notar que a diferencial da entalpia é considerada função apenas na temperatura, isto é,  $dH = c_p dT$ . Usando agora a equação (1) desenvolvida anteriormente, teremos:

$$\sum_i (C_i V) c_{p,i} \frac{dT}{dt} + \sum_i H_i F_{i,in} - \sum_i H_i F_i + \sum_i H_i \sum_j v_{i,j} (-r_j V) = \dot{Q} - \dot{W}_S + \sum F_{i,in} H_{i,in} - \sum F_i H_i$$

Comutando a ordem do somatório duplo anterior:

$$\sum_i H_i \sum_j \nu_{i,j} (r_j V) = \sum_j (r_j V) \sum_i \nu_{i,j} H_i(T) = \sum_j (-r_j V) \Delta H_j^R(T)$$

temos a própria definição de calor de reação na temperatura  $T$ , da reação ' $j$ '. O balanço então pode ser escrito da seguinte forma:

$$\sum_i (C_i V) c_{p,i} \frac{dT}{dt} = \dot{Q} - \dot{W}_S - \sum_i F_{i,in} (H_i - H_{i,in}) - \sum_j (-r_j V) \Delta H_j^R(T)$$

Considerando, pela sua extensão, que apenas a reação de propagação tem efeito térmico apreciável, podemos escrever o balanço empregando a expressão adequada para a taxa, estendido para todas as espécies possíveis, bem como para todos os sítios considerados. O trabalho de eixo foi desprezado, bem como admitiu-se que se dá o resfriamento do reator através de uma camisa em temperatura constante; isto é mais realista se a vazão de fluido de resfriamento é alta.

$$(3) \quad \left( \sum_i C_i c_{p,i} \right) V \frac{dT}{dt} = -UA_t (T - T_c) - \sum_i F_{i,in} (H_i - H_{i,in}) - \sum_{j=0}^{ns-1} \Delta H_R^P(T) K_{P,j} M (\mu_{0,j}^L + \mu_{0,j}^{L,a}) V$$

O índice ' $j$ ' no somatório indica agora o tipo de sítio. A análise dimensional da equação acima nos traz que:

$$\left[ \frac{mol}{m^3} \right] \left[ \frac{J}{mol \times K} \right] \left[ m^3 \right] \left[ \frac{K}{s} \right] = \left[ \frac{J}{m^2 s K} \right] \left[ m^2 \right] \left[ K \right] - \left[ \frac{mol}{s} \right] \left[ \frac{J}{mol} \right] - \left[ \frac{mol}{m^3 s} \right] \left[ m^3 \right] \left[ \frac{J}{mol} \right]$$

O primeiro somatório representa a concentração total no sistema para um dado instante de tempo, ponderada pelo calor específico; isto nada mais é que a capacidade calorífica da mistura,  $\Phi = \Sigma(C_i c_{pi} V)$ , em J / K. Uma vez que as concentrações estão disponíveis durante a resolução conjunta com os balanços de massa, é possível, mediante os valores convenientes de calor específico em função da temperatura, determinar-se esta grandeza.

Com relação ao termo que envolve diferença de entalpias, vemos que o mesmo se refere às condições da entrada; portanto, a contribuição energética dos produtos da reação já foi considerada nos demais termos. Uma vez que a entalpia total da fase sólida é desprezável frente à de líquidos e gases, nas condições usuais, este somatório envolverá apenas três espécies: solvente, monômero e hidrogênio. A forma como serão tratadas as entalpias é mostrada adiante no texto.

Finalmente, resta o balanço térmico para a camisa de resfriamento. Aplicando as leis de conservação em grandezas mássicas para o fluido de resfriamento, temos o seguinte balanço simplificado:

$$\rho_c c_{p,c} V_c \frac{dT_{w,c}}{dt} = \rho_c c_{p,c} F_c (T_{c,in} - T_{w,c}) + UA_t (T - T_{w,c}) \quad (4)$$

onde o subscrito 'c' indica o fluido.

## 5. Cálculo das Propriedades do Polímero

### 5.1. Massa Moleculares

As massas moleculares médias do polímero serão dadas da maneira usual ( $\bar{M}$  é a massa molecular do monômero).

$$M_n = \bar{M} \frac{\mu_1^D + \sum_{j=0}^{ns-1} \mu_{1,j}^L}{\mu_0^D + \sum_{j=0}^{ns-1} \mu_{0,j}^L} \quad (\text{massa molecular média em número})$$

$$M_w = \bar{M} \frac{\mu_2^D + \sum_{j=0}^{ns-1} \mu_{2,j}^L}{\mu_1^D + \sum_{j=0}^{ns-1} \mu_{1,j}^L} \quad (\text{massa molecular média em peso})$$

$$PD = \frac{M_w}{M_n} \quad (\text{polidispersão})$$

## 5.2. Distribuição de Pesos Moleculares

A geração da curva de distribuição de pesos moleculares (DPM) será feita através da aproximação do grau de polimerização como uma variável contínua e a expansão das distribuições dos comprimentos de cadeia através de polinômios do grau de polimerização. Estes polinômios são ortogonais à uma distribuição estatística do tipo Schulz-Flory. As aproximações empregadas são as seguintes:

$$L_{r,j} \approx \alpha_L(j)^{(r-1)} \left[ L_{1,j} + \sum_{i=1}^{ng} w_i^L(j) \cdot (r-1)^i \right]$$

$$D_r \approx \alpha_D^{(r-1)} \left[ D_1 + \sum_{i=1}^{ng} w_i^D \cdot (r-1)^i \right]$$

Os coeficientes  $\alpha_{L,D}$  e  $w_i^{L,D}$  podem ser determinados substituindo-se as aproximações de  $L_{r,j}$  e  $D_r$  nas definições dos momentos estatísticos de ordem  $k$ :

$$\mu_{k,j}^L = \sum_{r=1}^{\infty} r^k L_{r,j} = \sum_{r=1}^{\infty} r^k \left\{ \alpha_L(j)^{(r-1)} \left[ L_{1,j} + \sum_{i=1}^{ng} w_i^L(j) \cdot (r-1)^i \right] \right\}$$

Após um pequeno desenvolvimento, e definindo a função somatório como

$$\Omega_{k,i}^L(j) = \sum_{r=1}^{\infty} r^k (r-1)^i \cdot \alpha_L(j)^{(r-1)} \quad \text{onde } 0 \leq k, \text{ e } i \leq ng.$$

Uma vez que para  $i = 0$  virá que

$$\Omega_{k,0}^L(j) = \sum_{r=1}^{\infty} r^k \cdot \alpha_L(j)^{(r-1)}$$

então:

$$\mu_{k,j}^L = L_{1,j} \cdot \Omega_{k,0}^L(j) + \sum_{i=1}^{ng} w_i^L(j) \cdot \Omega_{k,i}^L(j)$$

Resultado análogo pode ser obtido para os momentos  $\mu_k^D$  de ordem  $k$  do polímero morto. A relação acima, conhecidos os momentos através dos balanços de massa, gera o seguinte sistema de equações algébricas não lineares:

$$\begin{aligned} \mu_{0,j}^L &= L_{1,j} \cdot \Omega_{0,0}^L(j) + \sum_{i=1}^{ng} w_i^L(j) \cdot \Omega_{0,i}^L(j) \\ \mu_{1,j}^L &= L_{1,j} \cdot \Omega_{1,0}^L(j) + \sum_{i=1}^{ng} w_i^L(j) \cdot \Omega_{1,i}^L(j) \\ &\vdots \\ \mu_{ng,j}^L &= L_{1,j} \cdot \Omega_{ng,0}^L(j) + \sum_{i=1}^{ng} w_i^L(j) \cdot \Omega_{ng,i}^L(j) \end{aligned}$$

O sistema está completamente determinado ao achar-se as relações de recorrência entre os diversos somatórios. Isto pode ser feito empregando a definição dos mesmo, seguida de um mudança de índice e da expansão das séries resultantes:

$$\Omega_{0,0}^{L,D} = \frac{1}{1 - \alpha_{L,D}}$$

$$\Omega_{1,0}^{L,D} = \frac{\Omega_{0,0}^{L,D}}{1 - \alpha_{L,D}}$$

$$\Omega_{k,0}^{L,D} = \frac{\alpha_{L,D}}{1 - \alpha_{L,D}} \cdot \left[ \frac{\Omega_{0,0}^{L,D}}{\alpha_{L,D}} + \sum_{j=1}^{k-1} C_{k,j} \cdot \Omega_{j,0}^{L,D} \right] \quad \text{para } 1 \leq k \leq 2 \times ng$$

$$\Omega_{k,i}^{L,D} = \Omega_{(k+i),0}^{L,D} + \sum_{j=1}^{(i-1)} (-1)^j C_{i,j} \cdot \Omega_{(k+i-j),0}^{L,D} + (-1)^i \cdot \Omega_{k,0}^{L,D}$$

onde  $C_{k,j}$  e  $C_{i,j}$  são os coeficientes binomiais das expansões  $(r+1)^k$  e  $r^k (r-1)^i$  de, respectivamente:

$$C_{k,j} = \frac{k!}{j! \cdot (k-j)!} \quad e \quad C_{i,j} = \frac{i!}{j! \cdot (i-j)!}$$

Uma vez resolvido o sistema de equações acima, as frações ponderais ‘ $\varphi$ ’ dos pesos moleculares podem ser dadas através das seguintes relações:

$$\varphi_{L,j}(r) = \frac{r L_{r,j}}{\mu_{1,j}^L}$$

para o polímero vivo proveniente do sítio do tipo ‘ $j$ ’;

$$\varphi_L(r) = \frac{r L_{r,j}}{\sum_{j=0}^{ns-1} \mu_{1,j}^L}$$

para o polímero vivo total;

$$\varphi_D(r) = \frac{r D_r}{\mu_1^D}$$

para o polímero morto;

$$\varphi_P(r) = \frac{r(L_r + D_r)}{\mu_1^D + \sum_{j=0}^{ns-1} \mu_{1,j}^L}$$

para o polímero total.

### 5.3. Propriedades Físicas da Mistura

#### 5.2.1. Calor e Massa Específicas

Os calores específicos dos compostos que entram no reator são dados em função da temperatura, bem como as massas específicas. A contribuição do polímero para a massa específica da mistura depende do grau de cristalinidade do mesmo, que não é

calculada pela rotina. Desse modo, será empregada uma massa específica constante para o polímero.

### 5.2.2. Entalpias

A entalpia de uma fase vapor (mistura) pode ser calculada como segue:

$$h_v = h_m^{res} + \sum_n y_i h_{v,i}^{Gid}, \text{ onde o primeiro termo é a entalpia residual de mistura,}$$

obtida a partir de uma equação de estado apropriada, e  $h_{v,i}^{id} = h_{f,i}^{o,Gid} + \int_{298,15}^T C_{p,i}^{Gid} dT$  é a entalpia do gás ideal na temperatura  $T$  do sistema, calculada a partir da entalpia de formação do composto puro como gás ideal a 298.15 K e da integral do calor específico a pressão constante. A entalpia de um vapor puro é dada analogamente, mediante um termo de correção do desvio em relação à idealidade.

No cálculo da entalpia da fase líquida, serão negligenciadas neste trabalho as interações moleculares nesta fase, de modo que não consideraremos efeitos térmicos de mistura (solução é ideal). A entalpia de um líquido pode ser dada por

$$h_L = \sum_n x_i h_{L,i}^{id}, \text{ onde } h_{L,i}^{id} = h_{v,i}^{Gid} + h_{vsat,i}^{res} - \Delta h_i^{vap} \text{ é a entalpia do líquido ideal,}$$

composta da entalpia do gás ideal mais a entalpia residual do vapor saturado, descontada da entalpia de vaporização (calculado pela equação de Clapeyron-Clausius).

O calor específico de gás ideal é dado em termos de uma expansão na temperatura com coeficientes tabelados. Uma expressão deste tipo pode ser usada para o cálculo da entalpia da fase líquida:

$$h_L = \Delta H_f^{liq.puro}(289,15K) + \int_{298,15}^T c_p^{liq} dT$$

Onde  $\Delta H_f$  é a entalpia de formação do líquido puro a 298,15K. Estas expressões, porém, são menos precisas, seus parâmetros são mais raros de se encontrar em literatura; porém, são bastante mais práticas. Como o termo que envolve as entalpias é uma diferença, não haveria problema quanto ao estado de referência, pois ele é

cancelado, bastando computar a integral de  $c_p$  ao longo da elevação de temperatura no reator:

$$\sum_i F_i (H_{i,s} - H_i) = \sum_i F_i \int_{298.15}^T c_{p,i} dT$$

para o líquido ou para o gás. O uso deste método ou do cálculo termodinâmico mais rigoroso será ditado pela conveniência.

Onde for necessário neste trabalho, será empregada a equação de pressão de vapor de Wagner.

## 6. Propriedades Termodinâmicas e Equilíbrio de Fases

### 6.1. Volume da Fase Líquida

O volume molar da mistura líquida pode ser obtido pela relação de Hankinson-Thomson-Brost (conhecida por CoStaLD) que é superior em precisão à correlação de Rackett:

$$\frac{v_m}{v^*} = v_R^{(0)} (1 - \omega_{SRK} v_R^{(\delta)}) \quad \text{onde } v_m \text{ é o volume molar da mistura líquida, } v^* \text{ é}$$

um volume característico (tabelado),  $\omega_{SRK}$  é o fator acêntrico obtido com a equação de Soave-Redlich-Kwong, e  $v_R^{(0)}$  e  $v_R^{(d)}$  são funções apenas da temperatura reduzida:

$$v_R^{(0)} = 1 + a(1 - T_R)^{1/3} + b(1 - T_R)^{2/3} + c(1 - T_R) + d(1 - T_R)^{4/3}$$

$$v_R^{(\delta)} = \frac{(e + fT_R + gT_R^2 + hT_R^3)}{(T_R - 1.00001)}$$

onde  $0.25 < T_R < 1.0$ ,  $T_R = T / T_{c,m}$  é a temperatura reduzida de mistura, e  $a, b, \dots$ , são constantes tabeladas. Para a mistura, a regra de cálculo das constantes pseudo-críticas é a seguinte:

$$T_{c,m} = \frac{\sum_i \sum_j x_i x_j (v_{ij}^* T_{c,ij})}{v_m^*} \quad \text{para o temperatura crítica de mistura,}$$

onde  $x$  representa a composição, e

$$v_m^* = \frac{1}{4} \left\{ \sum_i x_i v_i^* + 3 \left( \sum_i x_i v_i^{*2/3} \right) \left( \sum_i x_i v_i^{*2/3} \right) \right\} \quad \text{e} \quad v_{ij}^* T_{c,ij} = \left( v_i^* T_{c,i} v_j^* T_{c,j} \right)^{1/2}$$

O fator acêntrico de mistura é dado por  $\omega_m = \sum_i x_i \omega_{SRK}$ .

Para um dado número de moles conhecido da fase líquida (negligenciando o volume ocupado pelo catalisador e considerando pequeno o volume de hidrogênio dissolvido), então, o volume de líquido estará determinado por  $V_L = v_m (M^L + E^L)$ , a menos do contribuição do polímero. Isto também determina o volume ocupado pela fase vapor:  $V^V = V^R - V^L$ . Com o número de moles da fase vapor, pode-se chegar ao volume molar deste, que determina a pressão da fase vapor.

## 6.2. Equilíbrio Termodinâmico

O cálculo termodinâmico considerará que:

i) o hidrogênio está na condição de gás supercrítico (seu ponto crítico é definido por  $T_c = 33$  K,  $P_c = 12.9$  bar). Nesse caso, empregar-se-á um modelo apropriado para relacionar  $H_2$  (quantidade de hidrogênio total) com  $H_2^L$  (hidrogênio dissolvido); isto é feito mediante a solubilidade do gás no solvente.

ii) o solvente (uma mistura de compostos homólogos do hexano) e o monômero, podem se distribuir ou não entre as fases líquida e vapor, dependendo, para isto, da resolução de um problema de 'flash' com temperatura e pressão especificadas pelo ponto de operação estacionário do reator.

De todo modo, supondo que as condições são tais que o equilíbrio termodinâmico possa ser alcançado, valerá

$\hat{f}_i^{g,v} = \hat{f}_i^L$  onde  $\hat{f}$  representa a fugacidade do componente  $i$ , que pode ser um gás (g), vapor (v), ou um líquido (L). O índice 'i' será referido como 1 para a mistura de

solventes, 2 para o propeno e 3 para o hidrogênio (aqui considerado soluto). Uma abordagem distinta é feita para o hidrogênio.

### 6.2.1. Monômero e Solvente

No cálculo do *flash* considerar-se-á o solvente como um composto único por simplicidade. A equação de estado de Peng-Robinson será empregada para o cálculo das propriedades referentes ao estado gasoso e o modelo de Wilson será empregado para a fase líquida. O método de cálculo das quantidades empregadas pode ser encontrado nas referências [12] e [13]. Então, o valor de  $K$  utilizado na seção 3.2. é dado por:

$$K_i = \frac{y_i}{x_i} = \frac{\gamma_i P_i^{sat} \phi_i^{sat}}{\hat{\phi}_i P}, \text{ onde}$$

$\gamma_i$  é o coeficiente de atividade da fase líquida, função da temperatura e da composição, obtido através do modelo de Wilson;

$P_i^{sat}$  é a pressão de vapor do composto (puro), obtido com a equação de Wagner;

$\phi_i^{sat}$  é o coeficiente de fugacidade do composto puro na pressão de saturação, obtido através da equação de estado, na pressão de saturação do composto;

$\hat{\phi}_i$  é o coeficiente de fugacidade do componente na mistura, obtido através da equação de estado, na pressão do sistema;

$P$  é a pressão do sistema.

Também os termos de não idealidade nas entalpias (parcelas residual na fase gás e de excesso na fase líquida) são dadas a partir dos modelos para cada fase: Peng-Robinson e Wilson, respectivamente. A entalpia do vapor saturado é dada através da equação de Clapeyron-Clausius. As entalpias ideais são dadas como função apenas da temperatura, conforme descrito na seção 5.2.2.

### 6.2.2. O Hidrogênio

O hidrogênio gasoso será solubilizado no líquido, uma mistura de monômero com solvente, valendo ainda a seguinte relação de equilíbrio de fases:

$\hat{f}_3^g = \hat{f}_3^L = \gamma_3 x_3 f_3^{\circ(L)}$  onde  $\gamma$  é o coeficiente de atividade do soluto,  $x$ , sua fração molar, e  $f^\circ$  a fugacidade no estado padrão (líquido) na temperatura do problema. Notar que este estado é hipotético, porém, é possível fazer-se uma extrapolação para este estado, como será mostrado adiante. Desse modo, então,

$$x_3 = \frac{\hat{f}_3^g}{f_3^{\circ(L)} \gamma_3}$$

A fugacidade do componente em mistura na fase gasosa pode ser dada por uma equação de estado. Devido aos testes experimentais apresentados em literatura (ver discussão abaixo no item ‘**Outros Comentários**’), utilizou-se para o cálculo desta fugacidade a equação de estado de Soave-Redlich-Kwong com uma regra de mistura de Huron-Vidal Modificada de Segunda ordem (MHV2). Este método emprega a Energia Livre de Gibbs mediante o modelo UNIFAC de modo a considerar as informações trazidas por este último em baixas pressões, bem como de seus parâmetros de interação. Isto permite uma melhor determinação dos parâmetros de mistura, o que se traduz numa melhor reprodução do comportamento da mesma. A equação de SRK neste caso é:

$$\frac{Pb}{RT} = \frac{1}{1-u} - \frac{\alpha}{u(u+1)}, \text{ onde } \alpha = \frac{a}{bRT} \text{ e } u = \frac{v}{b}$$

O parâmetro  $b$  é calculado através de uma regra de mistura linear:  $b = \sum_{i=1}^{nc} y_i b_i$ . O

parâmetro  $\alpha$  da mistura é obtido através da resolução da seguinte equação ( $nc$  é o número de componentes):

$$q_1 \left( \alpha - \sum_{i=1}^{nc} z_i \alpha_i \right) + q_2 \left( \alpha^2 - \sum_{i=1}^{nc} z_i \alpha_i^2 \right) = \frac{g_{UNIFAC}^E}{RT} + \sum_{i=1}^{nc} y_i \ln \frac{b}{b_i}$$

Os parâmetros do modelo UNIFAC são dados, bem como  $q_1$  e  $q_2$  para a equação SRK,  $q(\alpha) = -\left(1 + \ln(u^0 - 1) + \alpha \ln \frac{u^0 + 1}{u^0}\right)$ , e  $u^0 = \frac{1}{2} \left[ (\alpha - 1) - \sqrt{\alpha^2 - 6\alpha + 1} \right]$  é a menor raiz que se obtém quando usa-se a condição de pressão nula na equação de estado. Espera-se que este procedimento estime bem a interação na fase vapor.

Então, pode-se usar o coeficiente de fugacidade no cálculo da fração solubilizada obtido com a regra de mistura otimizada:

$$\ln \hat{\phi}_i = \frac{b_i}{b} - \ln \left( z - \frac{bP}{RT} \right) + \frac{aP}{RT} \left( \frac{b_i}{b} - \frac{\partial a}{\partial T} \right) \ln \left[ \frac{z - bP/RT}{z} \right], \text{ onde } z = \frac{Pv}{RT} \text{ e } a \text{ e } b$$

são parâmetros obtidos para mistura. Empregando na expressão do equilíbrio termodinâmico:

$$x_3 = \frac{\hat{\phi}_3 P_3}{f_3^{\circ(L)}} \frac{1}{\gamma_3}$$

onde utilizou-se a pressão parcial de hidrogênio na fase gasosa ( $P_3$ ).

Na literatura também é recomendada uma relação que usa o Modelo de Solução Regular (Scatchard-Hildebrand) para o cálculo da solubilidade:

$$x_3 = \frac{\hat{f}_3^g}{f_3^{\circ(L)}} \exp \left( -\frac{v_3^L (\delta_3 - \bar{\delta})}{RT} \right), \text{ onde } \delta_3 \text{ é um parâmetro de solubilidade (e } \bar{\delta}$$

representa o equivalente para mistura de solventes,  $\bar{\delta} = \sum \Phi_i \delta_i$ ;  $\Phi$  é uma fr. volumétrica).

No entanto, devido à dificuldade da obtenção destes parâmetros, pode ser feito uso de outro modelo disponível, como o de Wilson, que, dentre os modelos de Energia Livre de Gibbs de Excesso é aquele que reproduz o comportamento experimental para o maior número de sistemas:

$$\frac{g^E}{RT} = -\sum_{i=1}^{nc} x_i \ln \left( \sum_{j=1}^{nc} x_j \Delta_{ij} \right) \text{ e } \ln \gamma_i = 1 - \ln \left( \sum_{j=1}^{nc} x_j \Delta_{ij} \right) - \sum_{j=1}^{nc} \frac{x_k \Delta_{ki}}{\sum_j x_j \Delta_{kj}}$$

Notar que ao substituir-se uma das relações acima na equação de equilíbrio, temos uma equação implícita em  $x_3$  que deve ser resolvida iterativamente (segundo a bibliografia consultada, a convergência é rápida). Ainda teríamos o problema da indeterminação das outras duas frações molares, o que pode ser eliminado se considerarmos que  $x_3 \ll 1$  e, portanto, a composição da fase líquida é substancialmente aquela calculada anteriormente.

#### 6.2.2.1. A Fugacidade no Estado Padrão

O valor de  $f_3^{\circ(L)}$ , a fugacidade do hidrogênio no estado padrão, é hipotético, e portanto, um método especial é necessário para o seu cálculo. Um destes métodos consiste na correção da fugacidade desde o estado estado de saturação até o estado atual:

$$f_3^{\circ(L)} = f_3^{S,L} \exp\left(v_3^L \frac{(P - P_3^S)}{RT}\right)$$

e o termo exponencial acima é conhecido como *fator de Poynting*. A fugacidade padrão no estado de saturação pode ser feita extrapolando-se a pressão de vapor do soluto supercrítico, obtida através de uma equação precisa, até a temperatura do sistema:

$$f_3^{S,L}(T_{sis}) = P_3^{sat}(T_{sis}) \phi_3^{sat}(T_{sis})$$

Outro método semi-empírico já bem estabelecido é o método de Chao-Seader, que se baseia em correlações baseadas nas pressões e temperaturas reduzidas para a predição do coeficiente de fugacidade no estado padrão. Para tanto, a relação de equilíbrio de fases é escrita na seguinte forma:

$$K_i = \frac{y_i}{x_i} = \frac{\gamma_i f_i^{\circ}}{\hat{\phi}_i P} = \frac{\gamma_i \hat{\phi}_i^{\circ}}{\hat{\phi}_i}$$

O valor de  $\hat{\phi}_i^{\circ}$ , coeficiente de fugacidade do líquido puro, é dado através de uma correlação baseada na Lei dos Estados Correspondentes:

$$\log \hat{\phi}_i^{\circ} = \log \hat{\phi}_i^{(0)} + \omega_i \log \hat{\phi}_i^{(1)}$$

Onde

$$\log \phi_i^{(0)} = A_0 + \frac{A_1}{T_r} + A_2 T_r + A_3 T_r^2 + A_4 T_r^3 + (A_5 + A_6 T_r + A_7 T_r^2) P_r + (A_8 + A_9 T_r) P_r^2 - \log P_r$$

$$\log \phi_i^{(1)} = -4,23893 + 8,65808 T_r - \frac{1,2206}{T_r} - 3,15224 T_r^3 - 0,025(P_r - 0,6)$$

O subscrito 'r' denota variável reduzida. Os parâmetros 'A' para o hidrogênio são dados na tabela a seguir, modificados por Grayson e Streed:

A <sub>0</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>
1,50709	2,74283	0,02110	0,00011	0	0,00858	0	0	0	0

Os valores acima não geraram resultados realistas nas simulações efetuadas. A modificação proposta por Jin, Greenkorn e Chao foi empregada:

$$\log \phi_i^{\circ} = A_0 + \frac{A_1}{T_r} + \frac{A_2}{T_r^2} + \frac{A_3}{T_r^3} + \left( A_4 + \frac{A_5}{T_r} + \frac{A_6}{T_r^2} + \frac{A_7}{T_r^3} \right) P_r + A_8 P_r^2 - \log P_r$$

A <sub>0</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>
0,86691	9,91959	-31,41416	37,4501	0,01837	-0,14604	0,85454	-1,16871	-6,91e-5

#### 6.2.2.2. Métodos Não-Preditivos

A solubilidade pode ser determinada a partir de métodos não-preditivos, isto é, métodos baseados na Lei de Henry:

$$H_{3,solv} = \gamma_3 f_3^{\circ}$$

Nesse caso, contudo, se faz necessário o conhecimento da constante de Henry para o sistema reacional específico, nas condições do sistema. Apesar de alguns métodos para correção de pressão (Eq. Krichevsky-Kasarnovsky) e temperatura estarem disponíveis, permanece a dependência da obtenção de valores tabelados. Este método, portanto, não será empregado.

## 6. Outros Comentários

### 6.1 Estratégia de Resolução

A solução rigorosa do problema dinâmico envolveria o acoplamento das relações de equilíbrio termodinâmico ao sistema de equações diferenciais ordinárias, para os componentes que se distribuem entre as fases. Isto significa transformar o sistema de equações diferenciais ordinárias num sistema algébrico-diferencial, onde as equações algébricas são altamente não lineares. Atualmente, o próprio Matlab dispõe de meios de resolver este tipo de problema, sendo a implementação presente possível; de qualquer modo, isto não elimina o fato de que se trata de um problema complicado e de maior esforço computacional.

O primeiro passo será o cálculo das composições de cada fase a fim de se obter a solução estacionária, que servirá como condição inicial da resolução dinâmica. Para a simulação dinâmica também será admitido que o equilíbrio termodinâmico é pouco deslocado pelo prosseguimento da polimerização, ou seja, que as constantes de partição entre fases não se alteram muito com a alteração das composições. No caso da simulação de distúrbios e alterações nas condições de operação que influem no equilíbrio de fases (nesse caso pressão e temperatura) as variáveis termodinâmicas serão reestimadas, conforme a necessidade. De todo modo, a fim de considerar o problema completo de *'flash'* com reação, será construída a rotina adequada para este caso.

### 6.2. A Questão Termodinâmica

O problema do equilíbrio envolvendo o hidrogênio não está bem resolvido ainda, de modo que não existe um consenso sobre a melhor maneira de considerar a sua solubilidade. Há duas alternativas possíveis: considerá-lo dentro de um *flash* junto com os demais componentes ou tratá-lo separadamente. No primeiro caso, temos problema de relacionar a fugacidade do gás supercrítico no estado padrão, enquanto que, no segundo, temos os inconvenientes de termo de recalcular as frações molares, além de desprezarmos os efeitos energéticos que o gás provoca sobre os demais constituintes da mistura. Uma vez que as relações de equilíbrio de fases, no entanto, são equivalentes,

neste trabalho será considerado o hidrogênio como participante do mesmo, diferindo apenas no método de estimação do valor de  $K$ .

O próprio modelo de solubilidade não é consensual, mas podemos ordená-los em grau crescente de complexidade:

a) Lei de Raoult:  $p_i = x_i P_i^{sat}$ , que negligencia não-idealidades nas fases (isto é, interações específicas) prevendo uma mesma solubilidade de 'i' em qualquer solvente;

b) Lei de Henry:  $p_i = k x_i = H_i x_i$ , que considera a interação soluto-solvente através de uma constante empírica que depende apenas da temperatura;

c) a equação de Krichevsky-Kasarnovsky e modificações:

$$\ln\left(\frac{f}{x_3}\right) = \ln H_{3,1}^{(P_1^S)} + \frac{\bar{v}_2^\infty (P - P_1^S)}{RT}$$

que também considera o efeito da pressão;

d) um modelo termodinâmico preditivo para o equilíbrio de fases como o mostrado no trabalho.

A equação de Krichevsky-Kasarnovsky foi empregada anteriormente na referência [1], e usualmente fornece bons resultados. No entanto, este modelo depende ainda de um parâmetro empírico específico do sistema, que pode não ser disponível, ou não apresentar a confiabilidade experimental desejada. Desse modo, a literatura consultada afirma que os métodos semi-empírico e preditivo nesse assunto competem entre si.

Uma vez que a proposição deste trabalho foi a de revisar a parte termodinâmica, na medida do possível se fará a comparação destes dois métodos. A modelagem termodinâmica considerada neste trabalho foi proposta a fim de representar um possível desvio com relação à Lei de Henry, dado que a concentração de hidrogênio é muito importante para o cálculo das propriedades finais do polímero. À primeira vista, pode soar estranho o uso de duas equações de estado diferentes num mesmo problema. O fato é que a equação de Peng-Robinson é a mais empregada em cálculos de equilíbrio de fases, apresentado excelentes resultados para hidrocarbonetos numa faixa de pressões baixas a moderadas. No entanto, a fim de investir-se mais na parte termodinâmica, resolveu-se fazer uso da regra de mistura proposta, uma vez que o trabalho em que é apresentada mostra vários teste feitos com a mesma, inclusive para o hidrogênio (e,

ainda, com a apresentação de parâmetros reestimados para as equações). No entanto, uma vez que os resultados mostrados no artigo eram específicos para a equação de Soave-Redlich-Kwong, os outros casos necessitando de um desenvolvimento adicional, preferiu-se adotar o modelo recomendado.

Um argumento extra para o uso da relação rigorosa de equilíbrio é que todos os modelos aqui apresentados constituem-se, em pior ou melhor grau, de aproximações da igualdade de fugacidades:  $f^V = f^L$ . Afirmar que  $f^V = p_i$  (pressão parcial) é negligenciar as interações da fase vapor. Uma vez que a maior parte dos dados experimentais para a constante de Henry provém de sistemas constituídos de gás mais soluto pouco volátil, é de se esperar algum desvio desta fugacidade. No entanto, como todo modelo, este grau de detalhamento pode ser acessório, estando os efeitos principais capturados por uma relação simples do tipo  $p = k x$ .

## 7. Bibliografia Consultada:

### Polimerização, Reatores e Modelagem:

1. Secchi, A. R. *Modelagem e Simulação Dinâmica de Produção de Polipropileno em Suspensão*, UFRGS-OPP, Abril/99
2. Bhagwat, M. S.; Bhagwat, S. S.; Sharma, M. M. *Mathematical Modeling of the Slurry Polymerization of Ethylene: Gas-Liquid Mass Transfer Limitations*. Ind. Eng. Chem. Res. **1994**, *33*, 2322-2330.
3. Gupta, S. K.; Sau, M. *Modelling of a semibatch polypropylene slurry reactor*. Polymer, **1993**, *34*, 4417-4426.
4. Woo, S. I.; Choi, H. K.; Kim, J. H.; Ko, Y. S. *Prediction of the Molecular Weight of Polyethylene Produced in a Semibatch Slurry Reactor by Computer Simulation*. Ind. Eng. Chem. Res, **1997**, *36*, 1337-1342.
5. Gupta, S. K.; Sarkar, P. *Modelling of propylene polymerization in an isothermal slurry reactor*. Polymer, **1991**, *32*, 2842-2852.
6. Gupta, S. K.; Sarkar, P. *Simulation of propylene polymerization: an efficient algorithm*. Polymer, **1992**, *33*, 1447-1485.
7. Brandrup, J.; Immergut, E. H. *Polymer Handbook*, 3<sup>a</sup> Ed. John Wiley & Sons, **1989**, New York.
8. Ray, W. H.; Debling, J. A.; Zacca, J. *Dynamic Modeling of Product Grade Transitions for Olefin Polymerization Processes*. AIChE Jr., **1994**, *40* n°3, 506-520.
9. Coni Jr., O. L. P.; *Modelagem e simulação dinâmica da polimerização via catalisadores Ziegler-Natta heterogêneos: Cálculo da distribuição de pesos moleculares*. Dissertação de Mestrado, COPPE/UFRJ, **1992**.
10. Fogler; H. S. *Elements of Chemical Reaction Engineering*. Prentice-Hall, **1992**, Englewood Cliffs.
11. Luyben, W. L. *Process Modeling, Simulation, and Control for Chemical Engineers*. McGraw-Hill Kogakusha, **1972**.
12. Manual do Polyred, UWPREL, Winsconsin, EUA, 1992.

**Termodinâmica:**

12. Walas, Stanley. *Phase Equilibria in Chemical Engineering*. Butterworth Publishers, **1985**, Boston.
13. Reid, R. C.; Prausnitz, J. M.; Poling, B. E. *The Properties of Gases and Liquids*. McGraw-Hill, **1987**, 4<sup>a</sup> Ed., New York.
14. Prausnitz, J. M. *Molecular Thermodynamics of Fluid-Phase Equilibria*. Prentice-Hall International, **1969**, Englewood Cliffs.
15. Gallant, R. *Physical Properties of Hydrocarbons*, Volume 1. Gulf Publishing Company, **1968**, Houston.
16. Dahl, S.; Fredenslund, A.; Rasmussen, P. *The MHV2 Model- A UNIFAC Based Model for Prediction of Gas Solubility and Vapor-Liquid Equilibria at Low and High Pressures*. IVC-SEP (Phase Equilibria and Separation Processes), **1990**, SEP 9023.
17. Hankinson, R.W.; Thomson, G. H. *Calculate liquid densities accurately*. Hydrocarbon Processing, **Set. 1979**, 29,277-283.
18. Jin, Z. L.; Greenkorn, R. A.; Chao, K. C. *Correlation of Vapor-Liquid Equilibrium Ratio of Hydrogen*. AIChE Journal, **Jun. 1995**, V.41 No. 6, 1602-1604.
19. Sebastian, H. M.; Lin, H. M.; Chao, K. C. *Correlation of Solubility of Hydrogen in Hydrocarbon Solvents*. AIChE Journal, **Jan. 1981**, V.27 No.1, 138-147.
20. Gmehling, J.; Li, J.; Schiller, M. *A Modified UNIFAC Model. 2. Present Parameter Matrix and Results for Different Thermodynamic Properties*. Ind. Eng. Chem. Res., **1993**, 32, 178-193.
21. Gmehling, J.; Lohmann, J.; Jakob, A.; Li, J.; Joh, R. *A Modified UNIFAC (Dortmund) Model. 3. Revision and Extension*. Ind. Eng. Chem. Res., **1998**, 37, 4876-4882.



## **2ª Parte: Implementação e Resultados**

## 1. Implementação

Todas as rotinas, correspondentes aos modelos desenvolvidos anteriormente, foram implementadas no ‘*software*’ MatLab, versões 4.2 e 5.3. A primeira versão é muito mais rápida e robusta na resolução de equações algébricas não-lineares, como as obtidas através dos modelos estacionários, e a versão 5.3. apresenta vantagens para a construção e manipulação de modelos dinâmicos.

Para efeitos de apresentação, os resultados serão listados em três tópicos: rotinas termodinâmicas, rotinas estacionárias e rotinas dinâmicas.

## 2. Rotinas Termodinâmicas

A validade das rotinas termodinâmicas foi testada no estado estacionário, através de cálculos de ‘*flash*’ com temperatura e pressão especificadas. Rotinas principais:

`vleshoot`: sua resolução fornece a estimativa inicial para o problema de equilíbrio de fases. Considera Lei de Raoult, isto é,  $K = y/x = P^{sat} / P$ ;

`vlereact`: arquivo de resolução estacionária do equilíbrio de fases, gerando as frações molares de cada fase, a fração vaporizada global e a entalpia resultante. Considera a contribuição não-ideal das entalpias, os coeficientes de atividade e fugacidade, a regra de mistura não-linear e o modelo de fugacidade padrão do hidrogênio.

Rotinas auxiliares:

`ddhvap`: entalpia de vaporização do componente puro (eq. Clapeyron-Clausius);

`hvapfeed`: o mesmo, mas com argumentos passados na chamada da função e não como `globals`;

`costald`: calcula volumes molares de líquidos saturados e misturas líquidas.

`Gienth`: calcula a entalpia do composto como gás ideal a uma dada temperatura;

`entfeed`: calcula a entalpia e o estado físico das correntes de alimentação do reator;

`entfug`: entalpia residual e coeficiente de fugacidade da mistura (eq. Peng-Robinson);

`entvap`: calcula a entalpia residual do composto puro como vapor saturado;  
`fh2st`: fugacidade do hidrogênio no estado padrão, segundo três métodos;  
`mhv2exp`: regra de mistura Huron-Vidal Modificada, forma explícita;  
`mhv2imp`: regra de mistura Huron-Vidal Modificada, forma implícita;  
`presvap`: pressão de vapor do composto a uma certa temperatura;  
`solvent`: permite configurar uma mistura de até três solventes entre alguns pré-definidos, fazendo uma média ponderada de suas propriedades termodinâmicas;  
`srkmhv2`: coeficiente de fugacidade do hidrogênio em mistura, usando eq. Soave-Redlich-Kwong e regra de mistura Huron-Vidal Modificada de segunda ordem;  
`UNIFAC`: coeficiente de atividade e entalpia de excesso da solução usando o modelo UNIFAC modificado (UNIFAC-Dortmund);  
`utils`: várias propriedades dos compostos puros no estado de saturação;  
`Wilson`: coeficiente de atividade e entalpia de excesso usando o modelo de Wilson;

### *Testes e Resultados das Rotinas*

Conforme já apontado, testes foram realizados com a abordagem termodinâmica como maneira de verificar a validade da mesma. São eles:

- 2.1- Teste com a regra de mistura não-linear;
- 2.2- Teste com o modelo de solubilidade para o hidrogênio;
- 2.3- Teste com o resultado do equilíbrio líquido-vapor (usando VLERACT, Hysys e VLEQUIL).

#### 2.1. Teste sobre o benefício do uso da regra de mistura não-linear

Considerando-se uma mistura gasosa constituída de propeno, hexano e hidrogênio com a composição molar de [0.6 0.3 0.1] teremos os seguintes resultados, para dois níveis de pressão e quatro temperaturas:

	Pressão: 5 atm		Pressão: 10 atm	
	Temp.: 300 K	Temp.: 320 K	Temp.: 330 K	Temp.: 340 K
A° (regra convencional)	0.1008	0.0870	0.01620	0.01512
A (regra MHV2)	0.0990	0.0853	0.01589	0.01483
diferença, %	1.9	1.9	1.9	1.9
ln( $\phi$ , h <sub>2</sub> )°	0.0895	0.0772	0.1622	0.1482

$\ln(\phi, h_2)$	0.0914	0.0756	0.1580	0.1445
------------------	--------	--------	--------	--------

\*  $A = a_{mist} P / (RT)^2$ , e 'a' obtido para cada regra de mistura.

Como podemos ver, o uso de uma regra de mistura mais sofisticada acrescenta uma pequena melhora (por hipótese) ao resultado.

## 2.2.- Teste com o modelo de solubilidade do hidrogênio

Três métodos para estimação da fugacidade do hidrogênio no estado padrão foram testados: solubilidade ideal, método de Chao-Seader e método de Zin-Chao.

a. Solubilidade ideal:  $\log(P^{sat}) = -34,4134 \cdot (1/T) + 1,4787$  (extrapolado);

$P^{sat}(50^\circ\text{C}) = 23,5610$  atm;  $P^{sat}(70^\circ\text{C}) = 23,90$  atm;

$\ln\phi^{sat}(50^\circ\text{C}, 23,56$  atm) = 0,0071;  $\ln\phi^{sat}(50^\circ\text{C}, 23,90$  atm) = 0,0073;

FP(50°C) = 1,0280; FP(70°C) = 1,0269; (Fator de Poynting)

$v_m^{sat} = 0,0538$  l/mol; P= 10 atm; (Volume molar de líquido sat.)

$\phi^{sat}(50^\circ\text{C}, 10$  atm) = 2,439;  $\phi^{sat}(70^\circ\text{C}, 10$  atm) = 2,472;

Para o cálculo de  $\ln\phi^{sat}$  foi empregada a equação de estado de Peng-Robinson e para cálculo de  $v_m^{sat}$  a correlação Costald.

b. Correlação de Chao-Seader:

$\ln\phi^{sat}(50^\circ\text{C}, 10$  atm) = 1,433,8;  $\ln\phi^{sat}(70^\circ\text{C}, 10$  atm) = 1,721,5;

c. Correlação Zin-Chao:

$\ln\phi^{sat}(50^\circ\text{C}, 10$  atm) = 3,9281;  $\ln\phi^{sat}(70^\circ\text{C}, 10$  atm) = 3,8625;

O método ideal é sabidamente uma estimativa grosseira. A correlação de Chao-Seader não mostrou resultados realistas, e a correlação de Zin-Chao foi escolhida.

## 2.3.- Testes Comparativos do Resultado do Equilíbrio Líquido-Vapor

Os resultados obtidos com os modelos descritos anteriormente no texto foram comparados com aqueles obtidos através do software comercial '*Hysys.Process*' e com a rotina VLEQUIL do '*Polyred*'. O caso sobre a qual foram aplicadas as rotinas **vlereact** e os referidos programas é a seguinte (dados da alimentação obtidos em reator real da OPP-Petroquímica):

Corrente	Constituinte	Vazão Molar (Kgmol/h)	Temperatura (K)	Pressão (atm)
1	Propeno	20	320	22
2	n-Hexano	40	343	22
3	Hidrogênio	0.5	340	15

Temperatura do reator: 337 K;

Pressão do reator: 10 atm.

Alguns resultados são dados comparativamente na tabela a seguir:

Variável	Hysys	Vlereact	VLEQUIL
Fr. Molar de Propeno no Líquido	0,3246	0,3242	0,2559
Fr. Molar de n-Hexano no Líquido	0,6722	0,6712	0,7416
Fr. Molar de Hidrogênio no Líquido	0,00313	0,00461	2,511e-3
Fr. Molar de Propeno no Vapor	0,6442	0,7027	0,6489
Fr. Molar de n-Hexano no Vapor	0,0752	0,0744	0,1011
Fr. Molar de Hidrogênio no Vapor	0,2806	0,2229	0,2500*
Fração Vaporizada	0,0184	0,0167	-
K Hidrogênio	89,6	48,4	-

\* Esta é uma variável de entrada da rotina VLEQUIL e foi estimada a partir dos resultados com as outras duas rotinas.

Adicionalmente, a rotina calcula a entalpia em (cal/mol) das correntes de entrada, o que pode ser dar uma idéia da qualidade do método de estimação de entalpia empregado:

**propeno:** Hysys: 1779,1; VleReact: 2831,1;

**n-hexano:** Hysys: -45.220,6; VleReact: -45.347,3;

**hidrogênio:** Hysys: 287,47; VleReact: 289,45.

### 3. Rotinas Estacionárias

A solução estacionária foi obtida de modo a ter-se um idéia de grandeza das variáveis envolvidas, como condição inicial a ser perturbada para a simulação dinâmica e para se testar o funcionamento das rotinas que descrevem os modelos.

Uma vez que a resolução do conjunto de equações algébricas não-lineares gerado pelo modelo estacionário é bastante pesada computacionalmente, e esta resolução é muito dependente da estimativa inicial disponível para a resolução, modelos simples foram construídos a fim de gerar estas estimativas.

### 3.1. Descrição dos Modelos Estacionários

Todos os arquivos foram escritos empregando-se as concentrações e momentos como variáveis. Uma particularidade do modelo cinético empregado neste trabalho é que as equações referentes aos momentos de polímero vivo e morto podem ser obtidos explicitamente, para o estado estacionário, uma vez determinado o momento de ordem zero da distribuição:

$$\mu_{k,j}^L = \frac{M \left( K_{I,j} S_j + K_{P,j} \sum_{i=1}^k C_{k,i} \mu_{k-i,j}^L \right) V^L}{(K_{T,j} H_2 + K_{X,j}) V^L + q_L}$$

similarmente para  $\mu_k^D$ . Isto permite que se resolva o sistema de equações apenas com os momentos de ordem mais baixa, facilitando assim a resolução do sistema de equações. Os momentos de ordem superior podem ser calculados posteriormente de maneira recursiva em outra rotina.

Os arquivos criados foram os seguintes:

`coni`: modelo simplificado, escrito de acordo com o trabalho de Coni Jr. (ref. [1]). Considera apenas a fase líquida, com volume especificado, temperatura constante e especificada, escalonamento com valores fixos (independente do problema) para as variáveis, um tipo de sítio ativo e apenas as cinéticas de propagação de transferência para hidrogênio. Variáveis: concentrações de monômero, solvente, hidrogênio, sítios, polímero vivo e morto com grau de polimerização unitário, momentos de polímero vivo e morto de ordens zero, um e dois.

`fsolvf`: Modelo completo, considerando as fases líquida e vapor, volume especificado para a fase líquida, massa específica da mistura constante, número de sítios

igual a 'NS'. Cinéticas de iniciação, propagação, transferência para hidrogênio, desativação espontânea de sítios e desativação por 'donor'. Variáveis: concentrações de monômero, solvente e hidrogênio nas fases líquida e vapor, concentração de 'donor' na fase líquida, temperaturas do reator e da camisa de resfriamento, concentrações de sítios atáticos e isotáticos por tipo, concentrações de polímero com grau de polimerização unitário vivo e morto por tipo de sítio, momentos de polímeros vivos e mortos, atáticos e isotáticos de ordem até NG. Total:  $5NS+2NS \times NG+12+NG$  variáveis;

`fsolvef2`: o mesmo que o anterior, exceto que são calculados apenas os momentos de ordem zero da distribuição de pesos moleculares de polímero vivo e morto atático e isotático. Escalonamento baseado na estimativa inicial. Número de variáveis:  $5NS+12$ ;

`fsolvef3`: semelhante a `fsolvef2`, exceto que considera-se a variação de massa específica entre a entrada e a saída do reator, mediante uma soma ponderada pelas frações mássicas dos componentes da mistura reacional.

`fsolvef4`: semelhante a `fsolvef2`, exceto que considera-se os momentos de polímero vivo e morto, atático e isotático, de ordens zero, um e dois. Número de variáveis:  $9NS+16$ .

### 3.2. Dados de Simulação

Os dados a seguir são necessários para as simulações estacionária e dinâmica, e estão definidos nos arquivos de dados `default_values.m` (dados de operação), `kinetic.m` (informações cinéticas) e `data.m` (informações termodinâmicas). As vazões, temperaturas e pressões de alimentação são referentes à unidade de polimerização 'slurry' de propeno da OPP Petroquímica S.A., as constantes cinéticas foram buscadas nas referências [1] e [2], os dados termodinâmicos nas diversas referências listadas, bancos de dados do programas Polyred, Hysys e Led Zeppelin, banco de dados 'on-line' NIST... Os parâmetros numéricos para a resolução são declaradas na própria rotina e foram determinados através de tentativa e erro.

Algumas Variáveis Importantes da Simulação:

Variável	Nome na rotina	Valor	Unidade
----------	----------------	-------	---------

Vazão de alimentação de monômero	FM, Fm	2380,9	mol/min
Vazão de alimentação de solvente	FE, Fe	1527,31	mol/min
Vazão de alimentação de hidrogênio	FH2, Fh2	1,238	mol/min
* As vazões são informadas no vetor	Fin = [ Fm Fe Fh2 ]		
Temp. da corrente de monômero	Tin_mon	320	K
Temp. da corrente de solvente	Tin_solv	340	K
Temp. da corrente de hidrogênio	Tin_h2	340	K
Temp. de operação do reator	T0	340	K
Press. da corrente de monômero	Pin_mon	22	atm
Press. da corrente de monômero	Pin_mon	22	atm
Press. da corrente de monômero	Pin_mon	15	atm
Pressão de Operação do reator	P0	9.8	atm
Área de troca térmica do reator*	At	14.5	m2
Temp. do fluido de resfriamento*	Tcin	298	K
Vazão do fluido de resfriamento*	Fc	1	m3/min
Volume total do reator*	Vr	8.000	L
Volume máximo da mistura*	V0	6.000	L
Vazão de alimentação de 'donor'*	Fn	5e-6	mol/min
Coef. global de troca térmica *	U	2	cal/m.min.K
			.
Vazão mássica de catalisador	Fcat	4,88	g/min

\*Valores estimados pelo autor.

As informações cinéticas são dadas na tabela a seguir:

Variável	Nome na	Valor	Unidade
Fração de sítios ativos isotáticos	fsi	0,96	-
Atividade do catalisador	ativ_cat	1e-4	mol_s/g_cat
Fatores Pré-Exponenciais:			
Iniciação	Ai	2,061315e10	L/mol.min
Propagação	Ap	2,061315e10	L/mol.min
Transferência para Hidrogênio	At	7,729931e8	L/mol.min
Desat. Sítios e Term. Espontânea	Ax	2,061315e6	L/min

Desativação por 'donor'	Ad	2,061315e10	L/mol.min
Energias de Ativação:			
Iniciação	Ei	1,3278561e4	cal/mol
Propagação	Ep	1,3278561e4	cal/mol
Transferência para Hidrogênio	Et	1,3278561e4	cal/mol
Desat. Sítios e Term. Espontânea	Ex	1,3278561e4	cal/mol
Desativação por 'donor'	Ed	1,3278561e4	cal/mol
Entalpia da reação de propagação	DHR	-17.500	cal/mol

As constantes cinéticas são informadas às rotinas através de dois vetores:

AI = [Ai; Ap; At; Ax; Ad] e EI = [Ei; Ep; Et; Ex; Ed].

O efeito da distribuição de sítios sobre as constantes (energia de ativação) é calculado através da função `polycin`:

`[A, E, zj, fj] = polycin(check, AI, EI, ns, sigma, opt1, opt2, z, opt3)`

Onde:

`check`: vetor de seis elementos (0 ou 1) que indica quais as reações consideradas (p.ex., a reação de ativação de sítios é marcada como um zero na primeira posição);

AI e EI: constantes cinéticas vistas acima;

ns: número de sítios ativos considerados;

sigma: desvio padrão para a energia de ativação;

opt1: permite aplicar a distribuição sobre os fatores pré-exponenciais;

opt2: permite o fornecimento pelo usuário de uma distribuição distinta da *normal* (usada como 'default' pela rotina);

opt3: permite considerar volumes de ativação.

As saídas da função são as matrizes `A_si` e `E_si`, `zj`, a distribuição reduzida de sítios para o valor de `ns` dado, e `fj`, a fração resultante de cada tipo de sítio.

Por brevidade, as informações termodinâmicas empregadas (parâmetros críticos, coeficientes da eq. de Wagner, das expansões de  $C_p$  e da correlação Costald, parâmetros binários de interação dos modelos, entalpias padrão de formação) não são transcritas aqui, mas podem ser encontradas nos arquivos de dados. Os parâmetros de

interação entre grupos do modelo UNIFAC se encontram no próprio arquivo UNIFAC.m.

### 3.3. Resultados da Simulação Estacionária

A simulação estacionária tem a seguinte sequência:

- Carregamento dos arquivos de dados e determinação da distribuição de sítios;
- resolução do problema de equilíbrio líquido-vapor ideal;
- resolução do problema de equilíbrio líquido-vapor rigoroso;
- resolução do modelo simples de polimerização slurry (coni);
- cálculo dos momentos de ordem superior;
- resolução do problema completo (fsolvf2 ou fsolvf4);
- cálculo dos momentos de ordem superior.

As tabelas a seguir reproduzem alguns dos resultados obtidos para coni e fsolvf4 para o caso base, isto é, para aqueles dados mostrados anteriormente.

Modelo: Coni		
Número de Variáveis: 12		
Resíduo na última iteração: 2,73e-5		
Variável	Valor	Unidade
Concentração de monômero (líquido)	3,2678	mol/L
Concentração de hidrogênio (líquido)	1,369e-4	mol/L
Concentração de sítios ativos	4,336e-10	mol/L
Conc. de pol. vivo com $r = 1$	4,337e-10	mol/L
Mom. ordem zero polímero vivo	1,745e-6	mol/L
Mom. ordem zero polímero vivo	6,910e-3	mol/L
Mom. ordem zero polímero vivo	54,93	mol/L
Variáveis determinadas externamente		
Peso molecular médio numeral	1,663e+5	gmol/g
Peso molecular médio ponderal	3,332e+5	gmol/g
Polidispersão	1,9998	

Modelo: Fsolvf4		
Número de Variáveis: 27		
Resíduo na última iteração: 1,806e-12		
Número de sítios (NS): 3		
Desv. padrão energia. ativ.: 600 cal/mol		
Variável	Valor	Unidade
Concentração de monômero (líquido)	3,260	mol/L
Concentração de hidrogênio (líquido)	1,366e-4	mol/L
Concentração de sítios ativos	sítio 1	0,204e-10
	sítio 2	2,346e-10
	sítio 3	2,865e-10
Mom. ord. zero pol. vivo isotat.	sítio 1	2,112e-7
	sítio 2	8,536e-7
	sítio 3	3,019e-7
Temperatura do reator	340,64	K
Variáveis determinadas externamente		
Peso molecular médio numeral	1,8754e5	gmol/g
Peso molecular médio ponderal	5,5577e5	gmol/g
Polidispersão	2,964	

O resultado completo pode ser obtido rodando-se a rotina com seus valores 'default'. A comparação entre o modelo simplificado e aquele completo mostra que este modelo completo não prevê dispersões muito elevadas para o conjunto de dados usado.

A tabela a seguir traz os resultados do 'Polyred' para o mesmo conjunto de dados:

Programa: Polyred		
Número de Equações: 28		
Tolerância: 1e-6		
Número de sítios (NS): 3		
Desv. padrão energia. ativ.: 600 cal/mol		
Variável	Valor	Unidade
Concentração de monômero (líquido)	4,9501	mol/L
Concentração de hidrogênio (líquido)		mol/L
Concentração de sítios ativos	sítio 1	0,8226e-10
	sítio 2	0,2478e-09
	sítio 3	0,8226e-10
Mom. ord. zero pol. vivo isotat.	sítio 1	0,8746e-10
	sítio 2	0,2624e-09
	sítio 3	0,8746e-10

Temperatura do reator	326,35	K
Peso molecular médio numeral	63.890	gmol/g
Peso molecular médio ponderal	127.740	gmol/g
Polidispersão	1,9993	

Então, uma série de teste foi feita variando-se os parâmetros cinéticos, a fim de verificar-se a influência dos mesmos sobre os resultados.

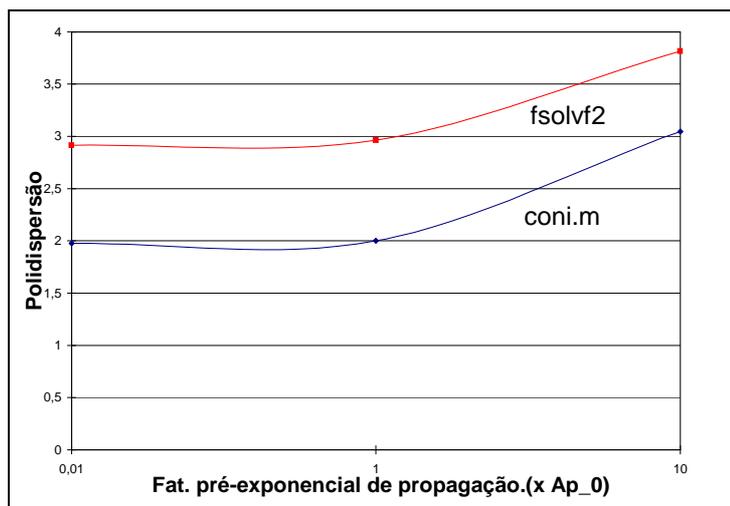


Figura 1. Efeito da variação da constante cinética de propagação sobre PD.

A tabela a seguir mostra o resultado da variação da constante de propagação em função do valor base ( $Ap_0$ ) sobre o peso moléculas médio numeral para os dois modelos:

(x $Ap_0$ )	0,01	1	10
<b>Coni</b>	1,711e+3	1,663e+5	1,096e+6
<b>Fsolvf2</b>	1,871e+3	1,875e+5	1,537e+6
<b>Polyred</b>	6,426e+3	6,389e+4	6,385e+5

Na figura 1 acima é mostrado o resultado desta variação sobre a constante cinética de propagação na polidispersão do polímero. Mantendo-se o valor de  $10 \cdot Ap_0$ , variou-se então o valor da constante de propagação da reação de transferência para o hidrogênio, também em função do valor base dado anteriormente  $At_0$ . O resultado é dado no gráfico a seguir:

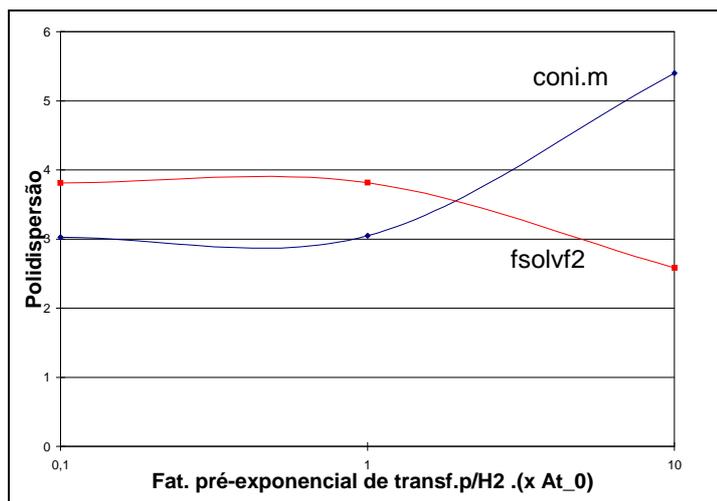


Figura 2. Efeito da variação da constante cinética de transferência sobre PD.

Similarmente ao caso anterior, vemos a influência da constante cinética de transferência sobre o peso molecular médio numeral do polímero para os dois modelos:

(x At_0)	0,1	1	10
<b>Coni</b>	1,110e+6	1,096e+6	5,393e+5
<b>Fsolvf2</b>	1,557e+6	1,537e+6	1,551e+6

Nota-se claramente que no modelo completo `fsolvf2` e variantes existe uma pequena influência da cinética de transferência para o hidrogênio, quer sobre a polidispersão, quer sobre o valor dos pesos moleculares médios. Isto deve se explicar porque, no modelo completo, há duas reações contribuindo para a diminuição do pesos moleculares: transf. para hidrogênio e terminação/ desativação, que foi preponderante neste caso.

Outros testes foram feitos com diversos parâmetros, mas estes só têm efeito sobre o modelo completo (não estão incluídas em `conim`). Foi tomado o valor de  $10 \cdot Ap_0$ , os demais parâmetros como nas tabelas de dados anteriores.

- reduzindo o fator pré-exponencial de desativação de sítios e terminação espontânea para 10% do valor base, temos um aumento de  $M_n$  para  $1,5365e6$  g/gmol, e um aumento da polidispersão para 5,058;

- o aumento em 50% do valor do desvio padrão da energia de ativação da distribuição de sítios acarretou um aumento da peso molecular médio numeral para 2,287e6 e da polidispersão para 4,2588;

- a diminuição do volume do reator e da fase líquida em 50% levaram a uma diminuição de PD para 3,2305 ( $M_n$  de 1,0625e6 g/gmol), e um aumento de 50% nas mesmas variáveis levou a 5,260 ( $M_n$  de 1,4679e6 g/gmol). Os volumes influenciam no tempo médio de residência da mistura, o que está relacionado à conversão.

- variações de  $\pm 25\%$  na solubilidade do hidrogênio não afetaram significativamente nem a polidispersão nem os pesos moleculares da solução do modelo completo, apenas a do modelo simplificado (gerando PD de 3,1 para a variação positiva);

- a aumento do número de sítios de 3 para 7 teve pequena influência na polidispersão (de 3,8 para 3,9). Isto está de acordo com as conclusões da referência [2].

#### 4. Rotina Dinâmica

Por comodidade, a condição inicial para a simulação dinâmica é dada através da solução estacionária, em torno da qual pode-se fazer as perturbações desejadas. Os resultados mostrados nesta seção tratam do caso no qual o equilíbrio de fases não varia imediatamente com as perturbações, mas é assumido uma alteração mais lenta deste em relação ao sistema.

O arquivo usado para a resolução dinâmica é `solvfile`, construído para a versão 5.3 do *Matlab*. Ele segue a nomenclatura básica já empregada anteriormente, e o equacionamento correspondente àquele mostrado na seção de modelagem deste trabalho.

A tabela a seguir traz a posição de cada variável no vetor de soluções ('estados'), no arquivo `solvfile`:

**NS**: número de sítios ativos de polimerização;

**NG**: ordem especificada para o polinômio interpolador da DPM;

\* Obs.: g.p. representa 'grau de polimerização'.

Variável	Descrição	Unidade
$n_M^L$	Número de moles de monômero na fase líquida	mol
$n_E^L$	Número de moles de solvente na fase líquida	mol

$n_{H_2}^L$	Número de moles de hidrogênio na fase líquida	mol
$S_j$	Concentração de sítios ativos isotáticos	mol/L
$S_j^a$	Concentração de sítios ativos atáticos	mol/L
$L_{1,j}$	Concentração de polímero vivo com g.p. = 1	mol/L
$D_1$	Concentração de polímero morto com g.p = 1	mol/L
$\mu_{0,j}^L$	Mom. de ordem 0 do polímero vivo isotático no sítio 'j'	mol/L
$\mu_{0,j}^{L,a}$	Mom. de ordem 0 do polímero vivo atático no sítio 'j'	mol/L
$\mu_{k,j}^L$	Mom. de ordem $k$ do polímero vivo isotático no sítio 'j'	mol/L
$\mu_{k,j}^{L,a}$	Mom. de ordem $k$ do polímero vivo atático no sítio 'j'	mol/L
$\mu_k^D$	Mom. de ordem $k$ do polímero morto	mol/L
$\mu_0^{D,a}$	Mom. de ordem zero do polímero morto	mol/L
$T$	Temperatura do reator	K
$T_{wc}$	Temperatura do fluido de resfriamento	K
$n_M^V$	Número de moles de monômero na fase vapor	mol
$n_E^V$	Número de moles de solvente na fase vapor	mol
$n_{H_2}^V$	Número de moles de hidrogênio na fase vapor	mol
$n_N$	Número de moles de 'donor'	mol

<b>Variável</b>	<b>Nome na rotina</b>	<b>Posição</b>
$n_M^L$	nml	1
$n_E^L$	nel	2
$n_{H_2}^L$	nh2l	3
$S_j$	Sj	4 : NS+3
$S_j^a$	Sja	NS+4 : 2NS+3
$L_{1,j}$	L1j	2NS+4 : 3NS+3

$D_1$	D1	3NS+4
$\mu_{0,j}^L$	mi0Lj	3NS+5 : 4NS+4
$\mu_{0,j}^{L,a}$	mi0Lja	4NS+5 : 5NS+4
$\mu_{k,j}^L$	mikLj	5NS+5 : 5NS+NS*NG+4
$\mu_{k,j}^{L,a}$	mikLja	5NS+NS*NG+5: 5NS+2NS*NG+4
$\mu_k^D$	mikD	5NS+2NS*NG+5: 5NS+2NS*NG+5+NG
$\mu_0^{D,a}$	mi0Da	5NS+2NS*NG+6+NG
$T$	T	5NS+2NS*NG+7+NG
$T_{wc}$	Twc	5NS+2NS*NG+8+NG
$n_M^V$	nmv	5NS+2NS*NG+9+NG
$n_E^V$	nev	5NS+2NS*NG+10+NG
$n_{H_2}^V$	nh2v	5NS+2NS*NG+11+NG
$n_N$	nn	5NS+2NS*NG+12+NG

As principais variáveis de entrada da simulação são:

Variável	Descrição	Nome	Tipo	Unidade
$A_{i,j}$	Fator pré-exponencial da reação do tipo 'i' no sítio do tipo 'j'	A	matriz $i \times j$	Idêntica à da taxa 'i'
$E_{i,j}$	Energia de ativação da reação do tipo 'i' no sítio do tipo 'j'	E	matriz $i \times j$	cal/mol
$F_{in}$	Vazão molar de alimentação (de monômero, solvente e hidrogênio)	F <sub>in</sub>	vetor $1 \times 3$	mol/min
$F_S$	Vazão molar total de sítios ativos (isotáticos e atáticos)	F <sub>S</sub>	Vetor $2 \times 1$	mol/min
$U$	Coefficiente global de troca térmica	U	Escalar	cal/m <sup>2</sup> .min
$A_r$	Área total de troca térmica do reator	A <sub>t</sub>	Escalar	m <sup>2</sup>
$V_r$	Volume total do reator	V <sub>r</sub>	Escalar	L
$T_{c,in}$	Temperatura de entrada do fluido de resfriamento	T <sub>c,in</sub>	Escalar	°C

$\Delta H_P^R$	Entalpia média da reação de propagação na faixa de temperaturas de operação	DHR	Escalar	cal/mol
$Oppoint$	Ponto normal de operação (P0, T0 e V0) do reator	Opp .	Vetor $3 \times 1$	atm, K, L
$F_C$	Vazão mássica de fluido de resfriamento	Fc	Escalar	Kg/min
$F_N$	Vazão molar de alimentação de 'donor'	Fn	Escalar	mol/min
$h_{feed}$	Entalpia total da alimentação	hf	Escalar	cal/mol
$x, y, \beta, Q$	Informações sobre o equilíbrio de fases ('flash'): composição e fração vaporizada	equil	Vetor $8 \times 1$	-
$Kp, q_{bias}$	Informações sobre os controladores (valores base de vazão e ganhos proporcionais)	Ctrl	Vetor $4 \times 1$	L/min. $\times$ (L ou atm)

As mesmas variáveis de entrada podem ser usadas para a geração de distúrbios, bastando para isso que se altere seu valor na chamada com relação ao valor estacionário.

Alguns testes foram feitos:

I. Aumento de 50% na vazão molar de monômero.

Os gráficos a seguir ilustram os resultados obtidos.

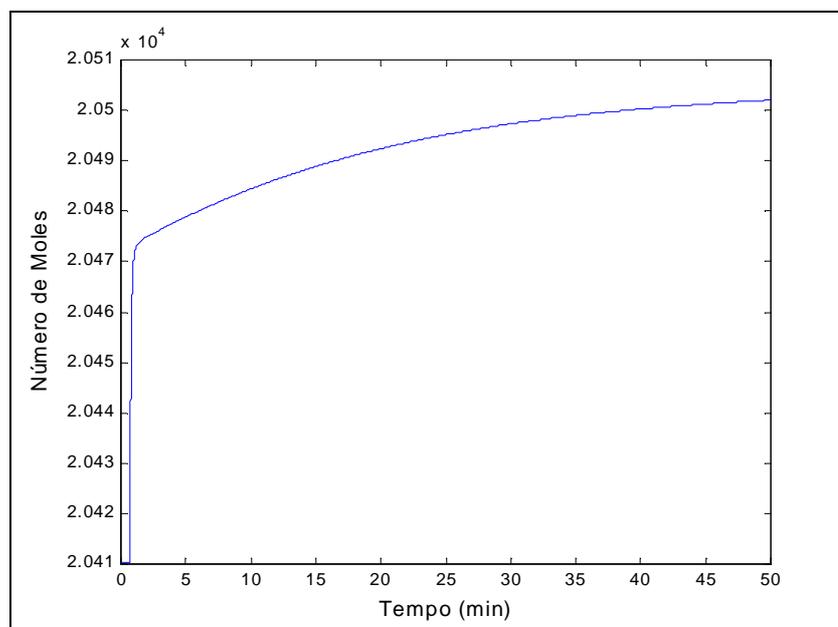


Figura 3. Resposta do número de moles de monômero na fase líquida frente a um aumento de 50% na vazão de entrada de monômero.

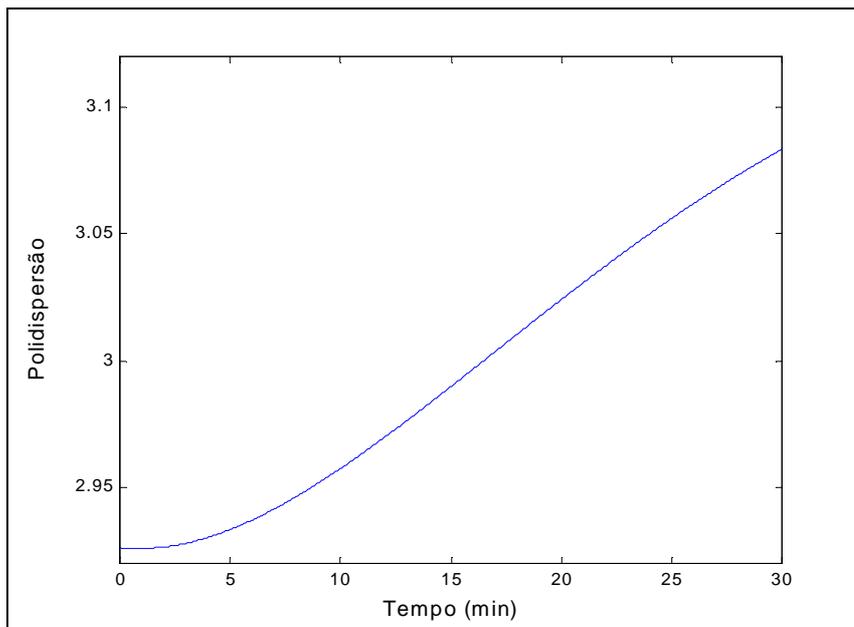


Figura 4. Resposta do índice de polidispersão frente a um aumento de 50% na vazão de entrada de monômero.

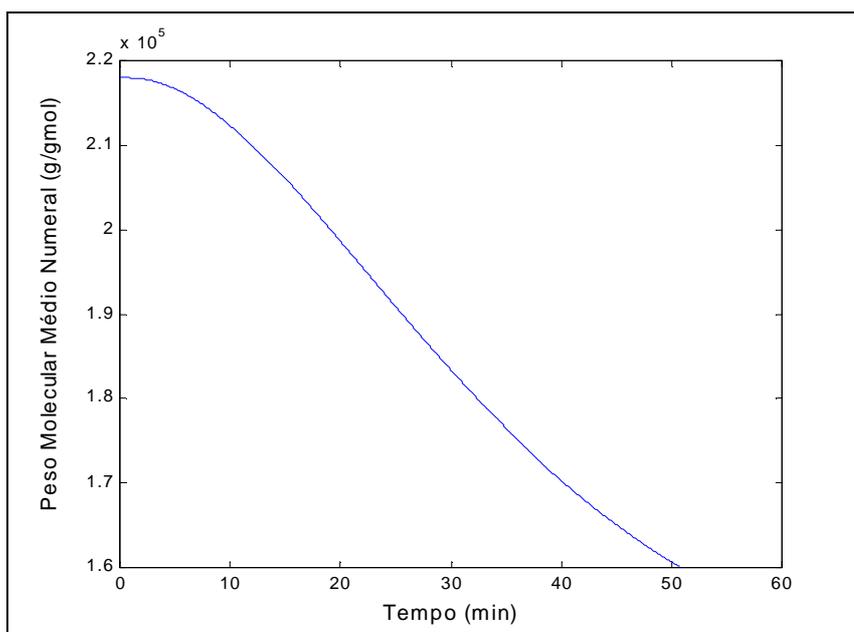


Figura 5. Resposta da peso molecular numéral frente à perturbação de 50% na vazão de entrada de monômero.

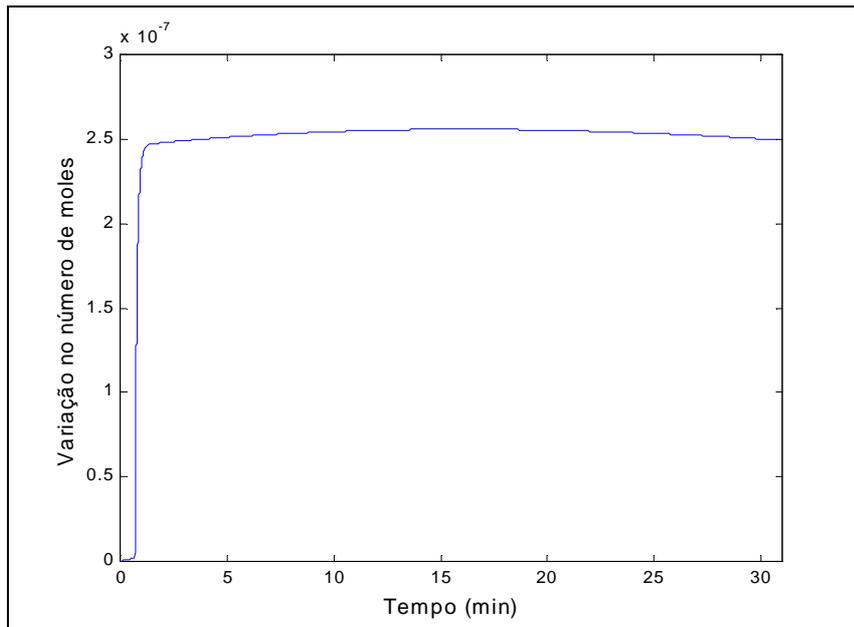


Figura 6. Variação do número de moles de hidrogênio na fase líquida para um aumento de 50% na vazão de entrada de hidrogênio.

\* Polidispersão estacionária após aumento de 50% na vazão molar de hidrogênio: 2,8873

## 5. Distribuição de Pesos Moleculares (DPM)

A distribuição de pesos moleculares, conforme descrito na parte deste trabalho referente à modelagem, é feita através da aproximação da quantidade de polímero com grau de polimerização ‘ $r$ ’ através de um polinômio que é função do grau de polimerização. A obtenção dos parâmetros que definem esta distribuição é feito através da resolução de um conjunto de equações não-lineares, em função dos momentos determinados através dos balanços materiais. Portanto, a obtenção da curva de distribuição de pesos moleculares, a qualquer tempo, é obtida pela resolução do sistema de equações que gera os parâmetros da curva de DPM em paralelo à integração do sistema de equações diferenciais.

Para o estado estacionário, temos uma simplificação, que pode ser vista a seguir:

$$\frac{d(L_{r,j}V)}{dt} = (K_{p,j}L_{r-1,j}M - K_{p,j}L_{r,j}M - K_{T,j}L_{r,j}H_2 - L_{r,j}K_{X,j}V)V - L_{r-1,j}q_s = 0$$

o que traz que:

$$L_{r,j} = \frac{(K_{p,j}MV)}{\left(\left(K_{p,j}M + K_{T,j}H_2 + K_{X,j}V\right)V + q_s\right)} L_{r-1,j}$$

ou seja,

$$L_{r,j} = \left[ \frac{(K_{p,j}MV)}{\left(\left(K_{p,j}M + K_{T,j}H_2 + K_{X,j}V\right)V + q_s\right)} \right]^{(r-1)} L_{1,j} = \alpha_j^{(r-1)} L_{1,j}$$

## Conclusões e Comentários

O modelo implementado, que se baseia principalmente na distribuição de sítios, prevê, com alguma facilidade, polidispersões entre 2 e 6, porém, não parece prever valores muito mais altos do que isso. Aparentemente, o uso de um modelo muito sofisticado para o hidrogênio não mostra vantagens. A sensibilidade aos parâmetros cinéticos se dá conforme o esperado.

Apesar de que bastantes testes tenham sido feitos com as rotinas estacionárias, poucos puderam ser feitos com a simulação dinâmica, em parte, porque as mesmas são muito lentas (quase em tempo real). Mais resultados poderão ser gerados, conforme se fizer necessário.

Pela dificuldade de acesso, não puderam ser disponibilizados resultados comparativos gerados pelo Polyred. Apesar de duas visitas à OPP, alguns problemas ocorreram durante as simulações, sobretudo no que se refere à manutenção de uma fase gás no reator. Nesse caso também, com mais uma visita pode-se obter resultados confiáveis.

Foram construídas para a construção da curva de DPM as funções `sumsolve` e `mwdist` no Matlab. Mais uma vez, por um questão de tempo, não puderam ser obtidas as condições iniciais necessárias para a resolução do sistema de equações algébricas não-lineares que determinam os coeficientes da distribuição.

Outra lacuna é a ausência da simulação dinâmica referente ao caso em que considera-se o equilíbrio termodinâmico variante no tempo. Para tanto, bastaria acrescentar-se ao arquivo `solvfile` as equações relativas ao equilíbrio de fases e o somatório das frações molares e modificar-se um pouco a rotina de chamada da função.

Foram iniciadas uma série de interfaces para orientar o uso das rotinas, incluindo o gerenciamento de parâmetros, variáveis de entrada, seleção de variáveis para plotagem, etc... Pelo tempo demandado, acabaram por restar incompletas, pois a prioridade era a geração de resultados; porém, pelo tempo já investido é justo que elas sejam finalizadas e entregues mais adiante.

Por fim, resta a conclusão de que, mais do que resultados, resta um série de possibilidades de se explorar as rotinas construídas. Por todos os motivos anteriormente externados, o tempo foi um fator limitante, ou por, outra, os objetivos propostos foram muito ambiciosos. Apesar de incompleto, o trabalho foi feito de forma dedicada e criteriosa, tomando realmente inúmeras horas de procura de bibliografia, de desenvolvimento das rotinas, de testes e correção dos problemas. Isto tudo, é claro, é revertido como benefício pessoal por aquele que se dedica; e tenho certeza de que me beneficiei muito com este trabalho.