



DEPARTAMENTO DE ENGENHARIA QUÍMICA  
FACULDADE DE CIÊNCIAS E TECNOLOGIA  
UNIVERSIDADE DE COIMBRA

# Controlo Predictivo Não-Linear de Processos Químicos

João Filipe Marques Gândara

Dissertação apresentada à  
FACULDADE DE CIÊNCIAS E TECNOLOGIA DA UNIVERSIDADE DE COIMBRA  
para obtenção do grau de  
DOUTOR EM ENGENHARIA QUÍMICA  
ESPECIALIDADE DE PROCESSOS QUÍMICOS

Coimbra 2012

Trabalho financiado pela Fundação para a Ciência e Tecnologia ao abrigo do programa PRAXIS XXI (PRAXIS/BD/9689/96).

# Resumo

Neste trabalho é abordada a aplicação de um algoritmo de optimização dinâmica a processos químicos. Mais especificamente, é considerado a utilização de algoritmos de controlo predictivo não-linear (NMPC — *nonlinear model predictive control*), baseados em métodos sequenciais, aplicados a processos não-lineares instáveis em ciclo aberto. Como ponto de partida, é utilizado um algoritmo de NMPC já existente, onde é implementada uma estratégia sequencial. As limitações sentidas na aplicação deste algoritmo ao problema de interesse são identificadas, sendo propostas algumas estratégias para as ultrapassar. É também abordado o estudo das principais características de modelos dinâmicos de processos químicos de média dimensão, tendo em vista conhecer a forma adequada de supervisão dos processos correspondentes. As estratégias consideradas são aplicadas a um problema com características e dimensões realistas, o processo Tennessee Eastman (TE).

Utilizando como ponto de partida o código disponibilizado por [Downs e Vogel \(1993\)](#), o processo TE é caracterizado recorrendo a diversas técnicas, como a determinação dos valores próprios do modelo linearizado e a análise de ganhos estacionários. São também identificadas as regiões viáveis de operação, dando particular atenção às suas extremidades, nomeadamente as correspondentes ao custo mínimo de operação para uma dada taxa de produção, e à maximização da capacidade absoluta de produção. Esta caracterização permite identificar vários aspectos relevantes deste processo, quer do ponto de vista da sua operação (nomeadamente considerando os aspectos de natureza económica), quer do ponto de vista da selecção do tipo de estrutura mais adequada para o seu controlo, em diversas localizações da sua região viável de operação.

Algumas características do processo TE levam a que não seja possível a aplicação do algoritmo de NMPC de referência, considerado inicialmente. Esta impossibilidade resulta da dificuldade, comum aos métodos sequenciais de optimização dinâmica, em lidar de forma adequada com processos instáveis em ciclo aberto. Desta forma, as dificuldades encontradas traduzem-se num mau condicionamento numérico dos problemas de optimização resultantes. Numa primeira fase, em que apenas é considerado o problema de controlo do reactor que integra o processo, estas dificuldades foram ultrapassadas pela utilização de uma estrutura de controlo em cascata. A apli-

cação desta estrutura de controlo ao processo resulta num problema de optimização cujo condicionamento numérico também já é suficiente para o uso, com sucesso, do algoritmo de NMPC considerado inicialmente. Os resultados obtidos, quando comparados com os apresentados por outros autores, permitem confirmar que é desejável a utilização de NMPC em processos com características semelhantes às encontradas no processo TE.

O carácter heurístico da estrutura de controlo utilizada levou a que fossem procuradas abordagens mais sistemáticas de aplicação de NMPC a processos instáveis em ciclo aberto, nomeadamente através da reformulação do algoritmo utilizado. Partindo de uma interpretação das consequências da utilização de uma estrutura de controlo em cascata anterior, é proposta uma estratégia que permite a aplicação directa de algoritmos sequenciais de NMPC a processos instáveis em ciclo aberto. As características mais robustas da nova formulação de controlo devem-se à forma como são representadas e manipuladas as equações do modelo e as suas variáveis. Neste âmbito, a utilização de decomposições numericamente robustas revela-se de extrema importância. Para aplicação desta formulação a processos instáveis em ciclo aberto, é também apresentada uma estratégia que permite mitigar os problemas resultantes da necessidade de integração das equações do modelo em ciclo aberto. Finalmente, a capacidade da formulação de controlo resultante destas contribuições para lidar com processos instáveis em ciclo aberto e problemas numericamente mal condicionados é demonstrada pela sua aplicação directa ao processo TE, sem necessidade do uso de camadas de controlo adicionais ou de outras alterações no processo.



# Abstract

The present work focus on the application of dynamical optimization methods to chemical processes. More specifically, the problem of nonlinear model predictive control (NMPC) with nonlinear, open-loop unstable processes, using sequential methods, is addressed. As a starting point, an existing NMPC algorithm, implementing a sequential method, is used. Its limitations are identified, and strategies to overcome these limitations are proposed. The issue of using dynamical models for process characterization is also addressed. The different strategies considered in this work are tested with the Tennessee Eastman (TE) process model, a realistic large dimension problem.

Using its model, the TE process is characterized through the use of a number of tools, including the calculus of the eigenvalues of the linearized model and the analysis of stationary gains. The viable regions of operation are also identified, giving particular attention to its extremes, namely the ones corresponding to the minimization of operating costs with specified production rates, and the absolute maximization of the production capacity. This characterization allows the identification of several relevant features of the process, from the viewpoints of its operation and of the selection of an adequate control structure.

Because of its properties, it is not possible to directly apply the previous NMPC formulation to the TE process, as a result of the difficulties (common to sequential methods) in dealing with open-loop unstable processes. These difficulties are, in part, due to the poor numerical conditioning of the resulting optimization problems. In a preliminary step, where only the reactor control problem is considered, these difficulties are overcome by introducing of a multilayer (cascade) control structure. The use of the same structure with the entire process results in an optimization problem with sufficiently good numerical conditioning for the successful solution of the NMPC problem. The results obtained in this manner, when compared to control results reported by other authors, show that the use of NMPC is desirable with processes that display similar features to the TE process.

The heuristic aspect of this solution led to the search for a more systematic approach for NMPC application to open-loop unstable processes, namely by reformulating the control algorithm which was initially used. Developing a base idea that originated

from the interpretation of the consequences of using a multilayer control structure, a reformulation of the previous NMPC is proposed. The improved features of this new NMPC formulation are mainly due to the way in which the model equations and variables are used. In this context, the use of numerical robust factorizations allows this formulation to deal with severely ill-conditioned problems. To overcome the difficulties due to the open-loop integration of the model equations, a strategy based on the use of predictive horizons with varying dimension is presented. The ability of the new formulation, resulting from these contributions, to deal with open-loop unstable processes and numerically ill-conditioned problems is demonstrated by its direct application to the TE process, without the need of additional control layers or further changes in the original process.

# Agradecimentos

Ao meu orientador, Professor Doutor Nuno Oliveira, pelo seu exemplo de seriedade e rigor, que me obrigou sempre a procurar soluções sólidas para os problemas que foram surgindo. Obrigado por todas as perguntas difíceis que me fez e por ter insistido nelas até surgir uma resposta. Obrigado por tudo o que me ensinou, sobre optimização dinâmica, controlo e sobre muitas outras coisas.

Encontro-me na situação de apresentar esta dissertação no Departamento de Engenharia Química onde fui aluno e onde, antes disso, passei muito dias de férias escolares. Devo, por isso, muito a muitas das pessoas que fazem, ou fizeram, parte deste Departamento. Obrigado a todos os meus professores que, no Chimico, tanto me ensinaram, dentro e fora das salas de aula e dos laboratórios. À Professora Doutora Margarida Figueiredo e à Professora Doutora Cristina Gaudêncio, obrigado por todo o apoio e encorajamento. Não posso deixar de aqui lembrar a memória do Professor Doutor José Almiro e Castro, por me ter convencido a seguir o curso de Engenharia Química e por ter contribuído para despertar em mim o gosto pela investigação.

Ao pessoal não docente do Departamento de Engenharia Química, obrigado por toda a simpatia, por toda a pronta ajuda, sempre que dela precisei, e por todos os momentos bem passados na sua companhia. Um agradecimento especial ao pessoal do Laboratório de Cálculo Automático (António Gândara, Manuel Borges e Adamo Caetano) por toda a pronta e eficaz ajuda com os problemas informáticos.

Ao Lino Santos, obrigado por toda a ajuda, com o modelo do reactor, com o `gnuplot`, com o `TEX`, e por todos os bons momentos passados a falar de outros assuntos igualmente interessantes.

A todos os meus colegas do Departamento de Ciência e Tecnologia Alimentar da Escola Superior Agrária de Coimbra, obrigado por todo o apoio, ajuda e, especialmente, pela amizade. Em particular obrigado à Aida Moreira da Silva, ao Carlos Dias Pereira, à Goreti Botelho, à Inês Seabra, ao Ivo Rodrigues, ao João Noronha, à Marta Henriques, à Regina Nabais, ao Rui Costa e ao Rui Saraiva. Obrigado também a todo o pessoal não docente, em particular à Paula Oliveira, ao Jorge Viegas e ao David Gomes. Obrigado ao Francisco Gil e ao Cristovão Silva por toda a ajuda dada no início da minha actividade docente.

A todos os que, durante a minha permanência, passaram pelo Laboratório de Sistemas, particularmente ao Pedro Lima, ao Nuno Maltez e ao Filipe Neves. Um agradecimento muito especial à Dulce Silva por toda a ajuda e amizade.

Aos meus pais, António e Alda, e à minha irmã, Teresa.

À Augusta, por ser minha esposa, e ao David, por ser meu filho.

*À Augusta e ao David*



# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objectivos e organização da tese . . . . .	3
<b>2</b>	<b>Formulação matemática de problemas de optimização dinâmica</b>	<b>7</b>
2.1	Modelos dinâmicos . . . . .	7
2.2	Formulação base . . . . .	10
2.3	Estratégias de solução . . . . .	12
2.3.1	Métodos indirectos . . . . .	12
2.3.2	Métodos directos . . . . .	16
2.4	Solução de DAE's e cálculo de sensibilidades . . . . .	25
2.5	Solução do problema de optimização . . . . .	30
2.5.1	Programação quadrática sucessiva . . . . .	30
2.5.2	Aproximações da Hessiana . . . . .	36
2.5.3	Métodos de globalização . . . . .	39
2.5.4	Regiões de confiança . . . . .	43
2.5.5	Métodos de ponto interior . . . . .	44
<b>3</b>	<b>Análise do processo Tennessee Eastman</b>	<b>49</b>
3.1	Descrição do processo . . . . .	49
3.1.1	Modelo do processo . . . . .	54
3.1.2	Objectivos de controlo . . . . .	56
3.2	Resultados publicados por outros autores . . . . .	58
3.3	O reactor do processo TE . . . . .	68
3.3.1	Modelo do reactor . . . . .	68
3.3.2	Resposta em ciclo aberto . . . . .	70
3.3.3	Ganhos estacionários . . . . .	71
3.3.4	Linearização do modelo do reactor . . . . .	74
3.4	Caracterização do processo . . . . .	75
3.4.1	Linearização do modelo do processo . . . . .	76
3.4.2	Condições óptimas de operação . . . . .	77
3.4.3	Regiões de operação viáveis . . . . .	83
3.4.4	Ganhos estacionários . . . . .	92

3.5	Conclusões	105
<b>4</b>	<b>Controlo predictivo do processo Tennessee Eastman</b>	<b>109</b>
4.1	Formulação base de controlo	109
4.1.1	Modelo discretizado	112
4.1.2	Solução do problema de optimização	117
4.1.3	Extensão para atrasos nas saídas	118
4.1.4	Acção integral	121
4.2	Controlo do reactor do processo TE	123
4.2.1	Estabilização com compensador interno	124
4.2.2	Resultados	126
4.3	Controlo do processo TE	130
4.3.1	Estrutura de controlo	130
4.3.2	Resultados	132
4.4	Conclusões	152
<b>5</b>	<b>Reformulação do algoritmo de controlo predictivo</b>	<b>155</b>
5.1	Motivação	156
5.1.1	Exemplos de teste	157
5.1.2	Condicionamento numérico da matriz dinâmica	158
5.2	Reformulação do problema quadrático	159
5.2.1	Representação explícita do modelo linearizado	160
5.2.2	Decomposição da solução	162
5.3	Adaptação do horizonte de previsão	168
5.4	Comparação com outras estratégias	170
5.5	Resultados de aplicação aos exemplos de teste	172
5.5.1	Condicionamento numérico	173
5.5.2	Permutações de variáveis	175
5.5.3	Trajectórias de controlo	177
5.6	Aplicação ao processo TE	179
5.6.1	Condicionamento numérico	180
5.6.2	Aplicação da estratégia de horizontes adaptáveis	180
5.6.3	Resultados de controlo	184
5.7	Conclusões	187
<b>6</b>	<b>Conclusões globais</b>	<b>189</b>
6.1	Perspectivas de trabalho futuro	191
	<b>Bibliografia</b>	<b>194</b>
<b>A</b>	<b>Comportamento das variáveis manipuláveis do processo TE</b>	<b>207</b>
A.1	Regime de operação 10G/90H	207
A.2	Regime de operação 50G/50H	212



A.3	Regime de operação 90G/10H . . . . .	216
<b>B</b>	<b>Ganhos estacionários do processo TE</b>	<b>221</b>
<b>C</b>	<b>Cálculo das sensibilidades</b>	<b>243</b>
C.1	Implementação computacional . . . . .	243
C.2	Eficiência do cálculo das sensibilidades . . . . .	245
C.2.1	Métodos para o cálculo das sensibilidades . . . . .	246
C.2.2	Solução de sistemas de equações lineares . . . . .	247
C.2.3	Eficiência das rotinas de álgebra linear . . . . .	250
C.2.4	Tolerâncias no cálculo das sensibilidades . . . . .	251
C.3	Comparação dos métodos de cálculo da sensibilidades . . . . .	254
C.4	Conclusões . . . . .	260



# Lista de Figuras

3.1	O processo Tennessee Eastman . . . . .	50
3.2	Zona do processo TE descrita pelo modelo do reactor. . . . .	68
3.3	Resposta do reactor a um aumento de $+5^{\circ}\text{C}$ na temperatura de entrada da água de arrefecimento . . . . .	70
3.4	Resposta do reactor a um corte na alimentação de $A$ . . . . .	71
3.5	Valores em estado estacionário da pressão, temperatura e nível do reactor, em função das variáveis manipuláveis. . . . .	72
3.6	Regiões de operação viáveis . . . . .	84
3.7	Posição da válvula de vapor do <i>stripper</i> em função do custo total e do caudal mássico de $G$ , no regime de operação $50G/50H$ . . . . .	85
3.8	Amplitude de variação da posição da válvula de vapor do <i>stripper</i> em função do custo total e do caudal mássico de $G$ , no regime de operação $50G/50H$ . . . . .	86
3.9	Restrições activas nas condições de operação correspondentes aos custos máximo e mínimo no modo de operação $10G/90H$ . . . . .	87
3.10	Restrições activas nas condições de operação correspondentes aos custos máximo e mínimo no modo de operação $50G/50H$ . . . . .	89
3.11	Restrições activas nas condições de operação correspondentes aos custos máximo e mínimo no modo de operação $90G/10H$ . . . . .	89
3.12	Variáveis medidas em função do caudal de alimentação de $D$ . . . . .	93
3.13	Variáveis medidas em função do caudal de alimentação de $E$ . . . . .	96
3.14	Variáveis medidas em função do caudal de vapor no <i>stripper</i> . . . . .	98
3.15	Variáveis medidas em função do caudal de água de arrefecimento do reactor . . . . .	101
3.16	Variáveis medidas em função do caudal de água de arrefecimento do separador . . . . .	103
4.1	Formulação de MPC com horizontes móveis. . . . .	110
4.2	Estrutura de controlo com MPC e compensador . . . . .	124
4.3	Variação da parte real do maior valor próprio, com o ganho proporcional. . . . .	125
4.4	Resposta do reactor em ciclo fechado a variações aleatórias nas cinéticas das reacções. . . . .	127

4.5	Resposta do reactor em ciclo fechado a variações aleatórias na temperatura de entrada da água de arrefecimento. . . . .	129
4.6	Perfis das variáveis medidas para uma perturbação em degrau de +5°C na temperatura de entrada da água de arrefecimento do reactor. . . . .	133
4.7	Perfis de controlo para uma perturbação em degrau de +5°C na temperatura de entrada da água de arrefecimento do reactor. . . . .	136
4.8	Perfis das variáveis medidas para uma variação aleatória das cinéticas das reacções. . . . .	137
4.9	Perfis de controlo para uma variação aleatória das cinéticas das reacções. . . . .	140
4.10	Perfis das variáveis medidas para uma diminuição de 15% do caudal do produto. . . . .	141
4.11	Perfis de controlo para uma diminuição de 15% do caudal do produto. . . . .	143
4.12	Perfis das variáveis medidas para uma diminuição de 2.2% do <i>set-point</i> da pressão do reactor. . . . .	144
4.13	Perfis de controlo para uma diminuição de 2.2% do <i>set-point</i> da pressão do reactor. . . . .	147
4.14	Perfis das variáveis medidas para a mudança de <i>set-point</i> da composição do produto. . . . .	148
4.15	Perfis de controlo para a mudança de <i>set-point</i> da composição do produto. . . . .	151
5.1	Número de condição da matriz dinâmica em função do tamanho do horizonte de previsão, para os sistemas LTI estável e instável. . . . .	159
5.2	Estrutura da matriz das equações linearizadas do modelo. . . . .	161
5.3	Estratégias de solução de problemas de optimização dinâmica. . . . .	170
5.4	Número de condição da Hessiana projectada do problema sem pivotagem em função do tamanho do horizonte de previsão, para os sistemas LTI estável e instável . . . . .	174
5.5	Número de condição da Hessiana projectada do problema com pivotagem em função do tamanho do horizonte de previsão, para os sistemas LTI estável e instável. . . . .	174
5.6	Estrutura da matriz $M_k$ do sistema LTI instável, com e sem pivotagem. . . . .	176
5.7	Estrutura das matrizes $Q$ e $R$ resultantes da factorização de $M_k$ com pivotagem. . . . .	176
5.8	Perfil de controlo para o problema do sistema LTI instável, com pivotagem. . . . .	177
5.9	Perfil de controlo para o problema do sistema LTI instável, sem pivotagem. . . . .	178
5.10	Perfis de controlo para o problema do sistema LTI instável, sem pivotagem, com horizontes de controlo $p = 28, 29, 30$ e $31$ . . . . .	178
5.11	Perfis da posição da válvula da água de arrefecimento do reactor . . . . .	182
5.12	Previsões da pressão do reactor . . . . .	183

5.13	Perfis das variáveis medidas para uma perturbação em degrau de $+5^{\circ}\text{C}$ na temperatura de entrada da água de arrefecimento do reactor. . . . .	185
5.14	Perfis das variáveis manipuláveis para uma perturbação em degrau de $+5^{\circ}\text{C}$ na temperatura de entrada da água de arrefecimento do reactor. . . . .	186
A.1	Valores das variáveis manipuláveis no interior da região viável para o regime de operação 10G/90H. . . . .	208
A.2	Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 10G/90H. . . . .	210
A.3	Valores das variáveis manipuláveis no interior da região viável para o regime de operação 50G/50H. . . . .	212
A.4	Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 50G/50H. . . . .	214
A.5	Valores das variáveis manipuláveis no interior da região viável para o regime de operação 90G/10H. . . . .	216
A.6	Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 90G/10H. . . . .	218
B.1	Variáveis medidas em função do caudal de alimentação de <i>D</i> . . . . .	222
B.2	Variáveis medidas em função do caudal de alimentação de <i>E</i> . . . . .	224
B.3	Variáveis medidas em função do caudal de alimentação de <i>A</i> . . . . .	226
B.4	Variáveis medidas em função do caudal de alimentação de <i>A</i> e <i>C</i> . . . . .	228
B.5	Variáveis medidas em função da posição da válvula de reciclo do compressor. . . . .	230
B.6	Variáveis medidas em função da posição da válvula da corrente de purga. . . . .	232
B.7	Variáveis medidas em função do caudal de vapor no <i>stripper</i> . . . . .	234
B.8	Variáveis medidas em função do caudal de água de arrefecimento do reactor. . . . .	236
B.9	Variáveis medidas em função do caudal de água de arrefecimento do condensador. . . . .	238
B.10	Variáveis medidas em função da velocidade de agitação do reactor. . . . .	240
C.1	Representação esquemática da implementação do algoritmo de controlo predictivo não-linear. . . . .	243
C.2	Distribuição do tempo de CPU necessário para obtenção dos perfis de controlo para o processo TE. . . . .	244
C.3	Distribuição do tempo de CPU necessário do processo TE, sem perturbações, utilizando o método simultâneo corrector (DASPK) . . . . .	247
C.4	Estrutura do Jacobiano do modelo do processo TE, incluindo o compensador, no estado estacionário correspondente ao modo base de operação. . . . .	248

C.5	Variações normalizadas do tempo de CPU, do número de avaliações das equações do modelo, do número de passos do integrador e do número de avaliações do Jacobiano para a integração do modelo e das sensibilidades do processo TE utilizando a DDASAC. . . . .	253
C.6	Variações normalizadas do tempo de CPU, do número de avaliações das equações do modelo, do número de passos do integrador e do número de avaliações do Jacobiano para a integração do modelo e das sensibilidades do processo TE utilizando a DASPCK (método alternado corrector). . . . .	254
C.7	Variações normalizadas do tempo de CPU, do número de avaliações das equações do modelo, do número de passos do integrador e do número de avaliações do Jacobiano para a integração do modelo e das sensibilidades do processo TE utilizando a DASPCK (método simultâneo corrector). . . . .	255
C.8	Razões entre os tempos de CPU necessários para a simulação utilizando o método alternado corrector e a biblioteca LINPACK e as rotinas MA48, e as bibliotecas ATLAS e BLAS . . . . .	257
C.9	Razões entre os tempos de CPU necessários para a simulação utilizando os métodos alternado corrector e simultâneo corrector . . . . .	258
C.10	Razões entre a DASPCK (método alternado corrector, LINPACK e ATLAS) e a DDASAC (método alternado directo) . . . . .	260

# Lista de Tabelas

3.1	Modos de operação do processo Tennessee Eastman. . . . .	51
3.2	Balanço mássico e energético do processo TE, correspondente à operação no caso base. . . . .	52
3.3	Variáveis do processo TE medidas continuamente. . . . .	53
3.4	Variáveis manipuláveis do processo TE. . . . .	54
3.5	Variáveis medidas descontinuamente no processo TE. . . . .	55
3.6	Limites normais de operação e de segurança do processo TE. . . . .	57
3.7	Variáveis do modelo do reactor. . . . .	69
3.8	Valores próprios da matriz de estado do modelo linearizado do reactor. . . . .	75
3.9	Valores próprios da matriz de estado do modelo linearizado do processo. . . . .	77
3.10	Condições de operação para caudal máximo de produto. . . . .	81
3.11	Condições de operação para caudal máximo de G no produto. . . . .	82
3.12	Condições de operação para custo mínimo. . . . .	83
3.13	Restrições activas nas condições de operação correspondentes aos custos mínimo e máximo no regime de operação 10G/90H. . . . .	88
3.14	Restrições activas nas condições de operação correspondentes aos custos mínimo e máximo no regime de operação 50G/50H. . . . .	90
3.15	Restrições activas nas condições de operação correspondentes aos custos mínimo e máximo no regime de operação 90G/10H. . . . .	91
4.1	Emparelhamentos entre variáveis capazes de estabilizar o modelo linearizado do reactor. . . . .	126
4.2	Número de condição da matriz Hessiana para as diferentes configurações de controlo. . . . .	131
5.1	Condição numérica da matriz dinâmica e da matriz Hessiana do problema de controlo dos sistemas LTI estável e instável. . . . .	158
5.2	Condição numérica na solução do problema de controlo do sistema LTI estável . . . . .	173
5.3	Condição numérica na solução do problema de controlo do sistema LTI instável . . . . .	173

5.4	Variação do número de condição da Hessiana projectada do problema de controlo do processo TE ( $Q_y = Q_u = I, m = 5$ ). . . . .	180
5.5	Variação do número de condição da Hessiana projectada do problema de controlo do processo TE ( $Q_y = I, Q_u = 0, m = 5$ ). . . . .	180
C.1	Tempo de CPU utilizado pelo módulo <i>Integ</i> , percentagem deste devido às sensibilidades, número de avaliações das equações do modelo, $F$ , e da matriz do Jacobiano, $J$ , necessários para a simulação de 0.1 h do processo TE, sem perturbações, utilizando como integrador as rotinas <i>DDASAC</i> e <i>DASPK</i> (método simultâneo corrector). . . . .	246
C.2	Tempo de CPU necessário para integração das equações do modelo e das sensibilidades na simulação de 0.1 h do processo TE utilizando diferentes métodos de solução de sistemas lineares, para diferentes métodos de cálculo das sensibilidades. . . . .	249
C.3	Tempo de CPU, em segundos, necessário para a integração das equações do modelo e das sensibilidades na simulação de 0.1 h do processo TE, sem perturbações, utilizando as rotinas de álgebra linear com e sem optimização, para diferentes métodos de cálculo das sensibilidades. . . .	251
C.4	Número de passos do integrador, de avaliações da matriz do Jacobiano e tempos de CPU utilizando diferentes estratégias de solução de sistemas lineares, necessários para a simulação do processo TE, utilizando o método alternado corrector no cálculo das sensibilidades ( <i>DASPK</i> ). . .	256
C.5	Número de passos do integrador, de avaliações da matriz do Jacobiano e tempos de CPU utilizando diferentes estratégias de solução de sistemas lineares, necessários para a simulação do processo TE, utilizando o método simultâneo corrector no cálculo das sensibilidades ( <i>DASPK</i> ). . .	258
C.6	Número de passos do integrador, de avaliações da matriz do Jacobiano e tempos de CPU utilizando diferentes estratégias de solução de sistemas lineares, necessários para a simulação do processo TE, utilizando o método alternado directo no cálculo das sensibilidades ( <i>DDASAC</i> ). . . .	259
C.7	Comparação entre o método alternado corrector, <i>DASPK</i> , e o método alternado directo, <i>DDASAC</i> . . . . .	259



# Capítulo 1

## Introdução

A otimização de processos químicos é uma área bastante vasta, contando actualmente com um número de aplicações muito extenso (Biegler, 2010). Entre outros factores, a sua utilização crescente é devida ao facto de, na generalidades dos problemas que envolvem processos químicos, ser possível identificar um elevado conjunto de aplicações possíveis destas técnicas, com elevado potencial de retorno. Entre estas é por vezes observável uma interacção complexa entre critérios de desempenho de naturezas diversas, como por exemplo os resultantes de considerações económicas e de desempenho. Os aspectos anteriores tornam extremamente útil o recurso a técnicas de otimização nas mais diversas fases do ciclo de vida dos processos químicos, envolvendo nomeadamente as fases de especificação (projecto) e de operação (supervisão) destes. À observação anterior acresce ainda a disponibilidade crescente de modelos matemáticos detalhados, provenientes das diferentes áreas da Engenharia Química, que permitem, de forma mais ou menos rigorosa, descrever estes processos. Deste modo, as soluções obtidas por aplicação de algoritmos de otimização a processos químicos podem resultar em benefícios significativos, quer sejam o aumento da eficiência, a melhoria da qualidade dos produtos ou o aumento do lucro (através da maximização das quantidades produzidas, ou através da minimização dos custos).

Também de uma forma muito genérica, as aplicações de otimização a processos químicos podem ser divididas em problemas de carácter estacionário e dinâmico. Exemplos de problemas estacionários são os relacionados com a síntese e o projecto de processos químicos, incluindo o *flowsheeting*, em que deve ser determinada a sequência e tipo de equipamentos a utilizar, e efectuado o dimensionamento destes equipamentos. Também podem ser incluídos nesta classe os problemas de determinação das condições óptimas de operação e os problemas de sequenciamento, ambos relacionados com a forma como os processos são operados. Neste problemas não é, por vezes, considerado o comportamento transiente das diversas etapas processuais.

A maior disponibilidade de ferramentas e modelos que permitem descrever o compor-

tamento dos processos químicos em estado transiente tem levado também, recentemente, a um aumento do interesse nos problemas de otimização dinâmica (Biegler, 2010). Nestes problemas, por exemplo, já não se pretende apenas conhecer as condições óptimas de operação, mas também a forma como essas condições devem ser atingidas. Uma característica desta classe de problema é a necessidade de utilizar modelos dinâmicos, que sejam capazes de descrever o comportamento do processo ao longo do tempo. Consequentemente, estes modelos contêm geralmente equações diferenciais, sendo o seu tratamento mais complexo que o necessário para as equações algébricas. Exemplos de problemas de otimização dinâmica são a estimativa de estados ou parâmetros, a detecção de falhas e o controlo predictivo, que será o objecto de estudo deste trabalho.

O controlo predictivo (MPC — *model predictive control*) teve origem na indústria no final da década de setenta, com o aparecimento dos primeiros computadores de processo com capacidade significativa de cálculo, tendo-se verificado desde então importantes desenvolvimentos na sua aplicação (Camacho e Bordons, 2007). O termo MPC não designa uma classe específica de métodos, sendo antes usado para identificar as *estratégias de controlo que utilizam, de forma explícita, um modelo do processo na obtenção de trajetórias que permitem otimizar um critério de desempenho ou uma função objectivo*, de natureza bastante genérica. De uma forma geral, este tipo de controladores utiliza o modelo do processo para prever o seu comportamento em função da trajetória de controlo escolhida. Estas últimas podem então ser seleccionadas, de forma sistemática, através da utilização de algoritmos de otimização adequados. O seu sucesso prático é sem dúvida atribuível a diversos factores, entre os quais ao facto de esta classe de métodos corresponder à forma mais geral de formular o problema de controlo no domínio temporal, permitindo lidar explicitamente com restrições operacionais (Camacho e Bordons, 2007).

Ainda que a generalidade dos processos químicos seja não-linear, muitas das aplicações industriais de MPC são baseadas na utilização de modelos lineares (Qin e Badgwell, 2003; Camacho e Bordons, 2007). Isto é sem dúvida devido à maior facilidade com que este tipo de modelos pode ser identificado a partir de dados operatórios dos processos. Estes modelos também permitem um desempenho satisfatório quando o processo opera habitualmente próximo de um estado de referência, para o qual o modelo foi obtido. Já quando se pretende operar o processo em condições mais diversificadas, resultantes por exemplo da mudança frequente das características dos produtos ou da sua procura, deve ser considerada a utilização de modelos não-lineares. Este modelos, frequentemente baseados em princípios fundamentais, permitem prever o comportamento do processo numa gama muito mais alargada de condições de operação e são fundamentais para captar a dinâmica de processos mais complexos. Da utilização destes modelos resulta o controlo predictivo não-linear (NMPC — *non-linear model predictive control*). Este tipo de algoritmo de controlo apresenta uma maior complexidade, devida ao facto de serem considerados problemas de optimi-

zação com restrições não-lineares (correspondentes às equações do modelo). Ainda que a aplicação deste tipo de controlador ao nível industrial seja ainda algo limitada (Camacho e Bordons, 2007), o que em parte se deverá a dificuldade na obtenção dos modelos necessários, este é um assunto merecedor de atenção, no qual se têm verificados bastantes progressos e desenvolvimentos assinaláveis.

Uma revisão mais extensa sobre as características e o funcionamento dos algoritmos MPC e NMPC, quer do ponto de vista da sua aplicação quer do ponto de vista das suas propriedades teóricas, pode ser encontrada em inúmeras referências. Entes estas destacam-se Qin e Badgwell (1997), Camacho e Bordons (1998), Morari e Lee (1999), Qin e Badgwell (2000), Mayne et al. (2000), Kwon e Han (2005), Bock et al. (2006), Tatjewski (2007) e Rawlings e Mayne (2009).

## 1.1 Objectivos e organização da tese

O objecto primordial deste trabalho consiste na aplicação de estratégias de controlo predictivo não-linear, baseadas em métodos sequenciais, a processos químicos. Sendo bem conhecidas as dificuldades que, de forma geral, surgem na aplicação directa desta classe de métodos a processos instáveis em ciclo aberto, o principal objectivo desta tese consiste na proposta de estratégias alternativas que permitam, de forma sistemática, garantir a possibilidade dessa aplicação. Adicionalmente, pretende-se ainda testar o desempenho dos algoritmos desenvolvidos com modelos de média complexidade, que possam ser representativos dos encontrados na prática industrial.

É apresentada, em seguida, uma breve descrição do conteúdo e organização deste relatório.

### **Capítulo 2: Formulação matemática de problemas de optimização dinâmica**

A complexidade dos problemas abordados implica o uso de um vasto conjunto de ferramentas matemáticas, que engloba a formulação e a resolução de problemas de optimização e a resolução de sistemas de equações diferenciais e algébricas. No Capítulo 2 é efectuada uma introdução abreviada a estes assuntos. Ainda que esta revisão possa parecer um pouco longa, entende-se o domínio adequado estes tópicos como fundamental para o enquadramento dos resultados apresentados posteriormente. Desta forma, optou-se por apresentar, logo de início, os conceitos fundamentais necessários ao desenvolvimento do trabalho realizado, deixando para depois a apresentação das contribuições resultantes desta tese.

### **Capítulo 3: Análise do processo Tennessee Eastman**

Para comparação das diferentes estratégias de NMPC consideradas neste trabalho, será utilizado o processo Tennessee Eastman (TE), proposto por [Downs e Vogel \(1993\)](#). As principais características deste processo são a sua forte não-linearidade, a instabilidade em ciclo aberto e a importante interação entre as diversas unidades que o constituem. Neste Capítulo é apresentada inicialmente uma descrição do processo TE, sendo depois discutidos os principais resultados publicados por outros autores, em diversas aplicações relacionadas com o controlo e a optimização deste processo.

São posteriormente apresentados os resultados obtidos relativos à caracterização do comportamento do reactor do processo, quando considerado isoladamente, e de todo o processo TE. Neste estudo são determinados os valores próprios da matriz de estado do modelo linearizado e analisados os ganhos estacionários deste. Na caracterização do processo são também identificadas as regiões de operação viáveis, com recurso à determinação das condições de operação correspondentes ao custo mínimo (para uma dada taxa de produção), e à produção máxima absoluta. São também evidenciadas as conclusões retiradas desta caracterização, do ponto de vista do comportamento analisado e da estrutura de controlo que, de acordo com os resultados apresentados, será mais adequada para lidar com este processo.

### **Capítulo 4: Controlo predictivo do processo Tennessee Eastman**

Neste Capítulo é inicialmente efectuada uma descrição do algoritmo de NMPC desenvolvido por [Oliveira \(1994\)](#), bem como de algumas extensões desenvolvidas e implementadas para a sua aplicação em processos com atrasos constantes nas saídas, e para a utilização de acção integral nestes casos. Este algoritmo utiliza uma abordagem sequencial aos problemas de optimização dinâmica. Inicialmente, é apenas considerado o problema de controlo do reactor do processo TE. Verifica-se que as características deste reactor inviabilizam a aplicação directa do algoritmo de NMPC original. Este problema é entretanto ultrapassado pela utilização de uma estrutura de controlo em cascata, sendo utilizado um compensador interno entre o processo e o controlador predictivo. O compensador usado é escolhido de forma a estabilizar o reactor, antes da utilização do algoritmo de controlo predictivo.

No caso do controlo global do processo TE, verifica-se novamente não ser possível a aplicação directa do algoritmo de NMPC original. Desta forma, foi testada uma tentativa de uso do mesmo compensador interno, desenvolvido anteriormente para o reactor. Neste caso, esta abordagem já não se mostra agora suficiente para estabilizar globalmente o processo originando, apesar de tudo, algumas melhorias no seu funcionamento. As melhorias registadas permitem neste caso mesmo a solução do problema de controlo predictivo para o processo global, com o algoritmo base. Este comportamento dever-se-à sobretudo à melhoria significativa observada no condicionamento numérico do problema, proporcionada pelo uso do compensador interno.

São também apresentados os resultados de controlo obtidos para algumas das perturbações e mudanças de *set-point* propostas por [Downs e Vogel \(1993\)](#). Os resultados obtidos confirmam que a utilização de NMPC é desejável em processos com as características do processo TE.

### **Capítulo 5: Reformulação do algoritmo de controlo predictivo**

A aplicação do algoritmo de NMPC proposto por [Oliveira \(1994\)](#) ao processo TE permite demonstrar algumas das limitações da aplicação de métodos sequenciais a processos instáveis em ciclo aberto. Para alargar de forma sistemática o uso destes métodos a uma classe mais abrangente de problemas, neste Capítulo são propostas alterações à anterior formulação de NMPC. Estas alterações são motivadas pela interpretação dos resultados obtidos com a utilização de uma estrutura de controlo adicional, mais especificamente, por análise das suas consequências nas alterações introduzidas na estrutura do modelo e no novo papel desempenhado pelas variáveis do sistema. São apresentadas estratégias que permitem ultrapassar o mau condicionamento numérico dos problemas de optimização e as falhas na integração das equações do modelo em ciclo aberto. No caso do mau condicionamento numérico, o problema é resolvido pela utilização de uma representação alternativa das equações do modelo, que permite a utilização de factorizações matriciais numericamente robustas ao nível da resolução do problema de optimização. No caso de potenciais falhas na integração das equações do modelo, é proposta uma estratégia de adaptação da dimensão dos horizontes de previsão que pode ajudar a mitigar este fenómeno.

A formulação de NMPC resultante da introdução destas estratégias na formulação anterior é directamente aplicada ao processo TE, tornando finalmente possível a solução do problema de controlo predictivos através de uma metodologia sequencial, sem a necessidade de introdução de camadas adicionais de controlo.

### **Capítulo 6: Conclusões globais**

São aqui apresentadas as principais conclusões retiradas a partir dos resultados apresentados ao longo de todo o trabalho, bem como identificadas as principais contribuições para os assuntos abordados. São também discutidos alguns aspectos merecedores de consideração no futuro, de forma a permitir uma maior generalização e aplicabilidade do algoritmo de controlo reformulado.



## Capítulo 2

# Formulação matemática de problemas de otimização dinâmica

Neste Capítulo são apresentados alguns conceitos fundamentais relativos à formulação e solução de problemas de otimização dinâmica. Inicialmente são descritos os diversos tipos de modelos utilizados nestes problemas, bem como algumas das suas características mais relevantes. É depois apresentada uma formulação genérica dos problemas de otimização dinâmica, e de alguns dos seus casos particulares. Segue-se uma descrição das diferentes estratégias de solução destes problemas e dos métodos utilizados para, neste âmbito, resolver as equações dos modelos e das suas sensibilidades. Finalmente, são apresentados diferentes métodos para a solução de problemas de otimização resultantes.

### 2.1 Modelos dinâmicos

Em todas as aplicações de otimização dinâmica é fundamental a disponibilidade de um modelo matemático do sistema físico cujo desempenho se pretende otimizar. As características desse modelo são fundamentais quer para o desempenho, quer para a selecção do procedimento de otimização a utilizar. Neste trabalho serão considerados sistemas descritos por modelos constituídos por equações diferenciais e algébricas (DAE's, *differential algebraic equations*). As DAE's surgem naturalmente na modelação matemática de processos químicos em estado transiente. A parte diferencial do modelo corresponde frequentemente às equações de balanços de massa, energia e momento, enquanto que a parte algébrica pode dever-se a condições de equilíbrio, ao cálculo de propriedades termodinâmicas, ou outras quantidades auxiliares (Pantelides et al., 1988). As DAE's podem ainda provir da discretização espacial de sistemas de equações diferenciais parciais.

As condições para existência de uma solução para estas equações, bem como os métodos numéricos a aplicar para a sua solução, dependem da sua estrutura, pelo que importa começar por uma breve descrição de algumas formas estruturais. As DAE's, na sua forma mais genérica, podem ser representadas como

$$0 = F(\dot{v}(t), v(t), u(t), w(t), \theta, t) \quad (2.1)$$

onde  $v \in \mathbb{R}^{n_v}$  são as variáveis de estado,  $u \in \mathbb{R}^{n_u}$  são as variáveis de entrada manipuláveis,  $w \in \mathbb{R}^{n_w}$  são as variáveis de entrada não manipuláveis, onde podem ser incluídas incertezas e perturbações desconhecidas, e  $\theta \in \mathbb{R}^{n_\theta}$  são parâmetros do modelo, que se consideram invariantes com o tempo. A representação (2.1) é designada como totalmente implícita.

Frequentemente, as variáveis de estado  $v$  podem ser separadas em variáveis algébricas,  $z \in \mathbb{R}^{n_z}$ , que apenas aparecem em (2.1) na forma algébrica, e em variáveis diferenciais,  $x \in \mathbb{R}^{n_x}$ , que aparecem em (2.1) tanto na forma algébrica como na forma diferencial. Quando esta separação é possível, as equações (2.1) podem ser escritas como:

$$0 = F(\dot{x}(t), x(t), z(t), u(t), w(t), \theta, t) \quad (2.2)$$

Se para além de ser possível a separação das variáveis de estado, também se verificar que  $\partial F / \partial \dot{x}$  é não singular, o modelo (2.2) pode ser representado numa forma semi-implícita (Brenan et al., 1989)

$$\dot{x}(t) = f(x(t), z(t), u(t), w(t), \theta, t) \quad (2.3a)$$

$$0 = g(x(t), z(t), u(t), w(t), \theta, t) \quad (2.3b)$$

em que também são separadas as equações diferenciais das equações algébricas do modelo. De notar que (2.3a) é um sistema de equações diferenciais ordinárias (ODE's, *ordinary differential equations*), pelo que os modelos nesta forma podem ser vistos como ODE's com restrições.

Uma das propriedades das DAE's que condiciona os métodos utilizados para a sua resolução é o índice diferencial. De forma geral, o índice diferencial de (2.1) é o número de diferenciações de  $F$  necessárias para que  $\dot{v}(t)$  possa ser resolvido em função das restantes variáveis (Ascher e Petzold, 1998). No caso de DAE's na forma semi-implícita (2.3), estas serão de índice um se e só se  $\partial g / \partial z$  for não singular na vizinhança da sua solução (Brenan et al., 1989; Petzold et al., 1997).

O estabelecimento de condições iniciais para as DAE's é outro aspecto em que estas diferem das ODE's, sobretudo quando o seu índice é igual ou superior a dois. Em DAE's de índice um o problema da inicialização é mais simples, sendo no entanto necessário garantir que as condições iniciais para a integração do modelo são consis-



tentes com as equações que o constituem. Quando o índice diferencial é igual a um é possível arbitrar valores iniciais para as variáveis de estado diferenciais ( $x(t_0)$ ), ou para as suas derivadas ( $\dot{x}(t_0)$ ), tal como no caso das ODE's. Os valores iniciais para as restantes variáveis terão que ser obtidos por resolução das equações do modelo. [Brown et al. \(1998\)](#) apresentam um método de inicialização para DAE's de índice um que não necessita de mais informação para além da necessária para a sua integração.

Para DAE's de índice igual ou superior a dois, o problema da inicialização torna-se mais complexo, devido à existência de condições adicionais, para além das impostas pelas equações do modelo. Para sistemas com índice  $p$ , estas condições são as  $(p - 1)$  derivadas em ordem ao tempo das condições explicitamente impostas por algumas equações do modelo ([Ascher e Petzold, 1998](#); [Pantelides et al., 1988](#)).

A modelação, em estado transiente, de processos químicos pode levar ao surgimento de DAE's de índice igual ou superior a dois. Segundo [Pantelides et al. \(1988\)](#), tal é frequente em modelos que envolvam suposições de equilíbrio termodinâmico, o que se deve à existência de equações algébricas que relacionam estados diferenciais, levando à redução do número de graus de liberdade do modelo. Em alguns casos, é relativamente simples a reformulação do modelo de modo a obter DAE's de índice um. Esta reformulação pode ser feita, por exemplo, relaxando as especificações de algumas quantidades físicas ([Pantelides et al., 1988](#)). Se tal não for viável, podem ser utilizadas técnicas automáticas para a redução do índice das DAE's a um, através da análise estrutural do modelo e diferenciação simbólica ou automática. Exemplos destas técnicas são as propostas por [Pantelides et al. \(1988\)](#) e [Unger et al. \(1995\)](#), que também podem ser utilizadas na obtenção da informação necessária para a inicialização de DAE's de índice superior a um.

No caso em que as DAE's são lineares, com coeficientes constantes ou variáveis, as condições que  $F$  em (2.2) tem que satisfazer para assegurar a existência e unicidade de uma solução  $x(t)$ , para um determinado perfil de variáveis manipuladas  $u(t)$ , estão bem estabelecidas e podem ser encontradas, por exemplo, em [Brenan et al. \(1989\)](#). A obtenção destas condições quando  $F$  é não-linear é mais complexa, e os resultados conhecidos estão limitados a condições suficientes locais. Alguns destes resultados são apresentados em [Brenan et al. \(1989\)](#), [Reich \(1992\)](#) e [Kunkel e Mehrmann \(1998\)](#).

No restante deste trabalho será assumido que as DAE's que constituem os modelos utilizados são de índice um e que a solução  $(x(t), z(t))$ , para um determinado  $u(t)$ , existe e é única.

É frequente, em processos químicos, algumas variáveis do modelo serem directamente medidas. Neste caso, o modelo (2.3) pode ser aumentado por

$$y(t) = h(x(t), z(t), u(t), \theta, t)$$

onde  $y \in \mathbb{R}^{n_y}$  são as variáveis medidas do processo. Esta extensão ao modelo não

é, de forma geral, necessária uma vez que as variáveis medidas podem ser incluídas nas variáveis algébricas. No entanto, em algumas situações é importante considerar explicitamente quais as variáveis que são medidas, como no caso de problema de controlo, sendo nestes casos útil considerar o modelo apresentado.

Por serem muito comuns, existe um grande número de aplicações que permite obter a solução destas equações, bem como calcular as suas derivadas de primeira ordem em relação aos parâmetros. Este assunto é mais extensamente discutido na secção 2.4.

## 2.2 Formulação base

Os problemas de optimização dinâmica podem ser, genericamente, formulados como:

$$\min_{u(t), \theta, t_F} \Psi(\bullet) = \Theta(x(t_F), z(t_F), t_F) + \int_{t_0}^{t_F} L(x(t), z(t), u(t), \theta, t) dt \quad (2.4a)$$

$$\text{s.a } 0 = F(\dot{x}(t), x(t), z(t), u(t), w(t), \theta, t), \quad t \in [t_0, t_F] \quad (2.4b)$$

$$0 = M(x(t_0), z(t_0), t_0) \quad (2.4c)$$

$$0 = N(x(t_F), z(t_F), t_F) \quad (2.4d)$$

$$0 \geq S(x(t), z(t), u(t), w(t), \theta, t), \quad t \in [t_0, t_F] \quad (2.4e)$$

Os instantes  $t_0$  e  $t_F$  correspondem, respectivamente, ao início e fim do intervalo de tempo em que se pretende optimizar o desempenho do processo. Nesta formulação, as equações do modelo são consideradas na forma semi explícita (2.3), podendo, no entanto, ser considerada qualquer uma das formas apresentadas. São consideradas restrições no instante inicial de operação ( $M$ ), no instante final de operação ( $N$ ), bem como restrições em todos os instantes intermédios ( $S$ ). Estas últimas são usualmente designadas como restrições de trajectória (*path constraints*), por limitarem as trajectórias possíveis entre os estados inicial e final. As funções  $\Theta$  e  $L$  são consideradas contínuas e diferenciáveis em  $[t_0, t_F]$  excepto, eventualmente, num número finito de pontos.

A função objectivo (2.4a) é usualmente designada como estando na forma de Bolza. É constituída pelo termo de Lagrange ( $L$ ), que depende das variáveis de estado e de entrada durante todo o intervalo de tempo considerado, e pelo termo de Mayer ( $\Theta$ ), que depende do instante final do intervalo, bem como das variáveis de estado nesse instante (Rao e Qiu, 1993). Dependendo do objectivo do problema, estes termos podem ser considerados isoladamente.

Consoante a função objectivo, do conjunto de restrições e das variáveis de decisão consideradas, a formulação (2.4) pode acomodar diferentes classes de problemas, tais como minimização do tempo total de operação, controlo predictivo de processos

e estimativa de estados ou parâmetros. A minimização do tempo total de operação é um problema de particular interesse na operação de processos descontínuos. Neste tipo de problemas a função objectivo toma simplesmente a forma

$$\Psi(\bullet) = t_F$$

constituindo as restrições (2.4d) as condições que o processo deve atingir para a operação pode ser terminada. Este problema foi extensamente estudado por [Silva \(2005\)](#).

Nos problemas de controlo predictivo o objectivo pode, genericamente, ser definido como manter as variáveis medidas o mais próximo possível de uma determinada referência, sem actuação excessiva sobre as variáveis de controlo. Assim, a função objectivo destes problemas, na sua forma contínua, pode assumir a forma

$$\Psi(\bullet) = \int_{t_0}^{t_F} \left[ (y(t) - y_{sp}(t))^T Q_y (y(t) - y_{sp}(t)) + (u(t) - u_r(t))^T Q_u (u(t) - u_r(t)) \right] dt \quad (2.5)$$

onde  $y_{sp} \in \mathbb{R}^{n_y}$  é a referência para as variáveis medidas, geralmente designada *set-point*, e  $u_r \in \mathbb{R}^{n_u}$  é o valor de referência para as variáveis de controlo. Nos problemas de estimativa de parâmetros, o que se pretende é obter os valores dos parâmetros do modelo que minimizam os desvios entre os valores observados e os previstos pelo modelo para as variáveis medidas, num determinado horizonte temporal passado. Neste caso, e considerando desvios quadráticos, a função objectivo pode tomar a forma

$$\Psi(\bullet) = \sum_i (y_{mod}(t_i) - y_{obs}(t_i))^T (y_{mod}(t_i) - y_{obs}(t_i))$$

onde  $y_{mod}$  e  $y_{obs} \in \mathbb{R}^{n_y}$  são os valores das variáveis medidas previstos pelo modelo e observados, respectivamente. Este problema também pode ser formulado de forma a serem obtidas estimativas das variáveis de estado do modelo a partir das observações feitas num determinado intervalo de tempo.

Em processos químicos é muito frequente as restrições de trajectória (2.4e) serem exclusivamente constituídas por limites aos valores que as variáveis do problema podem tomar, constituindo por isso este um caso de particular interesse. Nesta situação, estas restrições são geralmente designadas como *bound constraints* e, por exemplo, no caso das variáveis medidas tomam a forma

$$y^L \leq y(t) \leq y^U, \quad t \in [t_0, t_F]$$

onde  $y^L$  e  $y^U$  são, respectivamente, os valores mínimos e máximos que as variáveis podem assumir. Podem, obviamente, ser consideradas restrições semelhantes para as restantes variáveis do problema.

Também as restrições no instante inicial (2.4c) assumem frequentemente formas mais simples, sendo a mais comum

$$x(t_0) = x_0,$$

$$z(t_0) = z_0$$

em que são atribuídos valores às variáveis de estado nesse instante, que representam o estado do processo no instante inicial.

Por uma questão de simplificação da notação, daqui em diante não será explicitamente representada a dependência das variáveis do problema com o tempo, excepto quando tal se torne necessário para uma maior clareza do exposto, ou quando se pretenda representar um valor destas num instante de tempo específico.

## 2.3 Estratégias de solução

As estratégias de solução para os problemas de optimização aqui considerados podem ser divididas em duas categorias, indirectas e directas. Nas estratégias indirectas, a solução do problema é obtida pela resolução das suas condições de optimalidade. Nas estratégias directas o problema original, de natureza contínua e dimensão infinita, é transformado num problema descontínuo e de dimensão finita, que pode ser resolvido directamente como um problema de programação não-linear (Cervantes e Biegler, 2009). Uma breve perspectiva histórica das estratégias aqui descritas, bem como dos métodos numéricos utilizados para a obtenção da sua solução, pode ser encontrada em Sargent (2000).

### 2.3.1 Métodos indirectos

#### Análise variacional

Os métodos indirectos baseados na análise variacional obtêm uma solução para o problema de optimização resolvendo as suas condições necessárias de optimalidade de primeira ordem (Cervantes e Biegler, 2009).

Consideremos o problema (2.4) com os parâmetros do modelo ( $\theta$ ) e o tempo final ( $t_F$ ) constantes, ignorando as restrições de trajectória (2.4e). Este problema pode ser

transformado no problema de optimização sem restrições

$$\min_{u(t)} \bar{\Psi} = \Theta(x, z, t_F) + \zeta^T M(x(t_0), z(t_0), t_0) + v^T N(x(t_F), z(t_F), t_F) + \int_{t_0}^{t_F} L(x, z, u, \theta, t) + \lambda^T(t) F(x, z, u, \theta, t) dt$$

onde  $\zeta$ ,  $v$  e  $\lambda$  são vectores de multiplicadores de Lagrange, com dimensão adequada, sendo que  $\lambda$  é também designado vector das variáveis adjuntas do problema. As condições necessárias de primeira ordem são obtidas por anulação da primeira variação de  $\bar{\Psi}$  (Bryson e Ho, 1969). Definindo o Hamiltoniano como

$$H(t) = L(x, z, u, \theta, t) + \lambda^T(t) F(x, z, u, \theta, t)$$

as condições de primeira ordem para o problema em estudo são (Geering, 2007)

$$\left( \frac{\partial F}{\partial \dot{x}} \right)^T \dot{\lambda} = \frac{\partial H}{\partial x} \quad (2.6a)$$

$$0 = \frac{\partial H}{\partial z} \quad (2.6b)$$

$$0 = \frac{\partial H}{\partial u} \quad (2.6c)$$

$$0 = \frac{\partial H}{\partial \lambda} \quad (2.6d)$$

$$0 = \left( \frac{\partial M}{\partial x} \right)^T_{t=t_0} \zeta - \left( \frac{\partial F}{\partial \dot{x}} \right)^T_{t=t_0} \lambda(t_0) \quad (2.6e)$$

$$0 = \left( \frac{\partial M}{\partial z} \right)^T_{t=t_0} \zeta \quad (2.6f)$$

$$0 = \left( \frac{\partial \Theta}{\partial x} \right)^T_{t=t_F} + \left( \frac{\partial N}{\partial x} \right)^T_{t=t_F} v + \left( \frac{\partial F}{\partial \dot{x}} \right)^T_{t=t_F} \lambda(t_F) \quad (2.6g)$$

$$0 = \left( \frac{\partial \Theta}{\partial z} \right)^T_{t=t_F} + \left( \frac{\partial N}{\partial z} \right)^T_{t=t_F} v \quad (2.6h)$$

onde  $\zeta$  e  $v$  são os multiplicadores associados às restrições nos instantes inicial e final do horizonte, respectivamente. As condições (2.6a) a (2.6d) são as equações de Euler–Lagrange, constituindo as restantes condições fronteiras necessárias à obtenção de uma solução. Estas condições fronteira também são designadas condições de transversalidade.

A existência de restrições de trajectória activas durante a trajectória óptima de controlo leva a que a condição (2.6c) não possa ser sempre satisfeita. Nesta situação a obtenção da trajectória óptima de controlo utilizando as condições (2.6) implica o conhecimento *a priori* das restrições que estarão activas, o que pode nem sempre ser

fácil.

Quando no problema de controlo óptimo são consideradas restrições à trajectória do tipo (2.4e), as condições de optimalidade são dadas pelo princípio do mínimo de Pontryagin que, de forma compacta, pode ser enunciado como (Rao e Qiu, 1993):

**Definição 1 (Princípio do mínimo de Pontryagin)** *Se  $\mathcal{U}$  for a região admissível para a trajectórias de controlo, então a trajectória de controlo  $u^* \in \mathcal{U}$  que minimiza a função objectivo do problema de controlo óptimo também minimiza o seu Hamiltoniano. Então,*

$$H(x^*, u^*, \lambda^*, t) \leq H(x^*, u, \lambda^*, t), \quad \forall u \in \mathcal{U}$$

onde:

- $H(x^*, u^*, \lambda^*, t)$  é o Hamiltoniano correspondente à trajectória de controlo óptima;
- $H(x^*, u, \lambda^*, t)$  é o Hamiltoniano relativo a uma trajectória de controlo admissível;
- $x^*$  e  $\lambda^*$  designam, respectivamente, as trajectórias das variáveis de estado e dos multiplicadores de Lagrange correspondentes a  $u^*$ .

As condições de optimalidade necessárias são agora obtidas por aplicação da análise variacional ao problema de optimização obtido pelo princípio do mínimo de Pontryagin.

Para o problema (2.4), o Hamiltoniano é definido como

$$H(t) = L(x, z, u, \theta, t) + \lambda^T(t)F(\dot{x}, x, z, u, w, \theta, t) + \mu^T(t)S(x, z, u, w, \theta, t)$$

onde  $\mu$  são os multiplicadores de Lagrange das restrições de trajectória. As condições de optimalidade para este problema são semelhantes a (2.6), sendo necessário considerar a condição adicional (Srinivasan et al., 2003):

$$0 = \mu^T(t)S(x, z, u, w, \theta, t)$$

As condições de optimalidade resultantes da análise variacional resultam num sistema de equações diferenciais e algébricas às condições fronteira (TPBVP, *two point boundary value problem*). Tal resulta do facto de serem especificadas condições para os valores das variáveis quer no instante inicial, (2.6e) e (2.6f), quer no instante final, (2.6g) e (2.6h).

É possível a obtenção de soluções analíticas para este problema em casos específicos como, por exemplo, o do regulador quadrático linear. Na generalidade das situações tal não é possível, sendo necessário recorrer a soluções numéricas, constituindo as condições fronteira a principal dificuldade na sua obtenção (Støren e Hertzberg, 1995). Entre os métodos numéricos que podem ser utilizados encontram-se o *multiple shooting*, *single shooting*, *invariant embedding* (que transforma o TPBVP num pro-

blema às condições iniciais, mediante a suposição de uma estrutura para a solução) e a discretização (diferenças finitas e colocação em elementos finitos) (Støren e Hertzberg, 1995). Apesar da aplicação destes métodos não ser trivial, Støren e Hertzberg (1995) consideram que o principal obstáculo à resolução de problemas de optimização dinâmica por análise variacional é o grande volume de manipulações simbólicas necessário para a obtenção das condições de optimalidade, sendo que em muitos casos as equações do modelo não têm forma analítica, ou não são diferenciáveis.

De notar ainda que a solução óptima de controlo obtida pela análise variacional é uma solução em ciclo aberto, que consiste numa trajectória de controlo óptima que conduz o sistema de um estado inicial até um estado final, assumindo que em todos os instantes o sistema se mantém nessa trajectória (Bryson e Ho, 1969). Não há qualquer realimentação (*feedback*).

### Programação dinâmica

Outra abordagem à resolução de problemas de optimização dinâmica é a programação dinâmica, desenvolvida por Bellman nos finais da década de 1950. Esta abordagem baseia-se na aplicação do princípio de optimalidade, que pode ser enunciado como

**Definição 2 (Princípio de optimalidade)** *Quaisquer que sejam o estado inicial e as decisões iniciais, numa política óptima todas as restantes decisões devem constituir uma política óptima relativamente aos estados que resultam da primeira decisão (Bellman, 1957).*

Na derivação das condições de optimalidade resultantes do princípio de optimalidade iremos considerar que o modelo do processo é apenas constituído por equações diferenciais ordinárias, existindo apenas variáveis de estado diferenciais. Neste caso, o modelo do processo pode ser representado como

$$\dot{x} = f(x, u, w, \theta, t)$$

Nestas condições, e considerando os parâmetros do modelo constantes, a aplicação do princípio de optimalidade ao problema (2.4) resulta na equação de Hamilton–Jacobi–Bellman (Lewis e Syrmos, 1995; Srinivasan et al., 2003)

$$-\frac{\partial \Psi^*}{\partial t} = \min_{u(t)} \left[ L + \left( \frac{\partial \Psi^*}{\partial \mathbf{x}} \right)^T f + \mu^T S \right] \quad (2.7)$$

onde  $\Psi^*$  é o valor mínimo da função objectivo. Este problema é resolvido de forma regressiva, partindo do instante  $t_F$  onde é imposta a condição fronteira (Lewis e Syrmos, 1995)

$$\Psi^*(x(t_F), t_F) = \Theta(x(t_F), t_F)$$

na superfície  $N(x(t_F), t_F) = 0$ .

A solução de (2.7) podem ser obtida analiticamente em alguns casos, como o do regulador quadrático linear sem restrições de trajectória (Lewis e Syrmos, 1995). No caso geral de sistemas não-lineares, tal não é possível, sendo necessário recorrer a métodos numéricos para a obtenção de uma solução. Este procedimento não é trivial devido, em parte, à sua natureza diferencial parcial (Esposito, 2009).

A aplicação do princípio de optimalidade a DAE's é um problema de muito maior complexidade. Uma extensa discussão deste assunto pode ser encontrada em Sjöberg (2008).

### 2.3.2 Métodos directos

O problema (2.4) pode também ser resolvido transformando-o num problema de optimização com um número finito de graus de liberdade. Esta transformação pode ser conseguida recorrendo à discretização da variável tempo e à parametrização das variáveis do problema. Ao contrário dos métodos indirectos, nesta classe de métodos o problema de optimização é directamente resolvido, ainda que de forma aproximada, sendo por isso designados métodos directos. Entre estes métodos podem ser ainda distinguidas diversas abordagens, conforme são parametrizadas todas ou apenas parte das variáveis do problema (Cervantes e Biegler, 2009). Sem perda de generalidade, na discussão que se segue será considerado o problema da determinação de uma trajectória óptima de controlo no intervalo  $[t_0, t_F]$ .

#### Métodos sequenciais

No métodos sequenciais apenas as variáveis de controlo, ou manipuláveis, são parametrizadas, sendo estes também designados como métodos de parametrização das variáveis de controlo (*control vector parametrization*) (Cervantes e Biegler, 2009). Esta parametrização pode ser feita recorrendo a um conjunto de funções escalares base  $\phi_{i,j}(t)$ :

$$u_j(t) = \sum_{i=1}^{n_a} a_{i,j} \phi_{i,j}(t), \quad j = 1, \dots, v.$$

Uma vez que as funções  $\phi_{i,j}(t)$  são especificadas à partida, os coeficientes  $a_{i,j} \in \mathbb{R}$  passam a constituir as variáveis de decisão do problema. A mesma parametrização pode ser considerada em todo o horizonte, ou este pode ser dividido em  $N$  intervalos, sendo utilizadas diferentes parametrizações em cada um deles, o que permite a utilização global de aproximações mais simples. Frequentemente são considerados intervalos de tamanho constante, com um comprimento designado intervalo de amostragem,  $T$ ,



sendo as variáveis de controlo mantidas constantes em cada um desses intervalos

$$u(t) = u_k, \quad t \in [t_k, t_{k+1}[$$

com  $k = 0, \dots, N - 1$  e  $t_k = kT$ , o que corresponde a uma aproximação polinomial de ordem zero.

Dadas condições iniciais,  $x(t_0)$  e  $z(t_0)$  e uma trajectória de controlo

$$\tilde{u} = [\tilde{u}_0^T \quad \tilde{u}_1^T \quad \dots \quad \tilde{u}_{N-1}^T]^T$$

as equações do modelo (2.4b) podem ser resolvidas no horizonte  $[t_0, t_F]$ :

$$0 = F(\dot{x}(t), x(t), z(t), \tilde{u}(t), t)$$

A solução deste problema são trajectórias,  $\tilde{x}(t, \tilde{u})$  e  $\tilde{z}(t, \tilde{u})$  que são apenas função de  $\tilde{u}$ . Substituindo na função objectivo (2.4a), obtém-se:

$$\Psi(\tilde{u}) = \Theta(\tilde{x}(\tilde{u}, t_F), \tilde{z}(\tilde{u}, t_F), t_F) + \int_{t_0}^{t_F} L(\tilde{x}(\tilde{u}, t), \tilde{z}(\tilde{u}, t), \tilde{u}, t) dt$$

De forma semelhante, a restrição no instante final (2.4d) é escrita como

$$0 = N(\tilde{x}(\tilde{u}, t_F), z(\tilde{u}, t_F), t_F).$$

As restrições de trajectória (2.4e) podem ser incluídas no problema parametrizado recorrendo a diferentes estratégias (Feehery e Barton, 1998; Vassiliadis et al., 1994). Uma destas estratégias é a adição de um termo na função objectivo que penaliza a violação destas restrições. Outra alternativa é a escrita destas restrições apenas em alguns instantes do horizonte, que podem coincidir com os extremos dos intervalos de amostragem utilizados na parametrização das variáveis de controlo. Utilizando esta última abordagem, as restrições de trajectória são introduzidas no problema parametrizado como:

$$0 \geq S_i(\tilde{u}) = S(x(\tilde{u}, t_i), z(\tilde{u}, t_i), u, t_i), \quad t_i = t_0 + iT, \quad i = 0, \dots, N$$

Com esta abordagem apenas é garantida a satisfação destas restrições nos instantes considerados. Feehery e Barton (1998) propõem ainda outra estratégia em que as restrições de trajectória são impostas durante a solução das equações do modelo.

O problema não linear resultante da parametrização das variáveis de controlo é:

$$\begin{aligned} \min_{\tilde{u}} \quad & \Psi(\tilde{u}) = \Theta(\tilde{x}(\tilde{u}, t_F), \tilde{z}(\tilde{u}, t_F), t_F) + \int_{t_0}^{t_F} L(\tilde{x}(\tilde{u}, t), \tilde{z}(\tilde{u}, t), \tilde{u}, t) dt \\ \text{s.a} \quad & 0 = F(\dot{x}(t), x(t), z(t), \tilde{u}(t), t) \\ & 0 = M(x(t_0), z(t_0), t_0) \\ & 0 = N(\tilde{x}(\tilde{u}, t_F), \tilde{z}(\tilde{u}, t_F), t_F) \\ & 0 \geq S_i(\tilde{u}) = S(x(\tilde{u}, t_i), z(\tilde{u}, t_i), u, t_i), \quad , t_i = t_0 + iT, \quad i = 0, \dots, N \end{aligned}$$

O procedimento para resolução deste problema pode, de forma simplificada, ser escrito como:

1. estimativa inicial do do perfil de controlo;
2. solução do modelo e obtenção das derivadas de primeira ordem em torno do perfil de controlo;
3. iteração do algoritmo de resolução do NLP;
4. actualização do perfil de controlo;
5. repetição dos passos 2 a 4 até se se verificar convergência.

Este algoritmo permite verificar uma característica importante destes métodos que é o facto de, em cada iteração, serem resolvidas as equações do modelo em todo o horizonte. Daqui resulta a necessidade de que todas as trajectórias de controlo consideradas sejam viáveis, podendo contudo ser implementadas apenas parcialmente no caso de ser necessária a interrupção prematura do algoritmo. Verifica-se também que, em cada iteração, são necessárias pelo menos as derivadas de primeira ordem da função objectivo e das restrições. Esta informação pode ser obtida quer recorrendo a aproximações de diferenças finitas, quer através do cálculo das sensibilidades. A obtenção de informação de primeira ordem recorrendo ao cálculo de sensibilidades paramétricas será discutida mais detalhadamente na secção 2.4.

A utilização de diferenças finitas é o método conceptualmente mais simples, consistindo na aproximação de uma variação infinitesimal por uma variação finita em torno do ponto onde se pretende obter a derivada. Os erros cometidos nesta aproximação tanto se podem dever ao facto de ser utilizada uma aproximação demasiado grosseira, como a erros de cancelamento quando é utilizado uma variação muito pequena, devendo a variação a utilizar ser um compromisso entre ambas, de forma a minimizar o erro cometido (Gill et al., 1983). Apesar da sua simplicidade, a utilização de diferenças finitas implica, geralmente, um elevado número de avaliações da função que se pretende diferenciar, o que no problema em estudo resulta na necessidade de um elevado número de integrações das equações do modelo. Outra desvantagem

é a dificuldade em controlar a exactidão das aproximações obtidas (Gerds, 2005). A sua utilização é desaconselhada por alguns autores na resolução de problemas de optimização dinâmicos (Binder et al., 2001).

A principal limitação apontada aos métodos sequenciais é a sua falta de robustez, sobretudo quando aplicados a processos que contenham modos instáveis (Ascher e Petzold, 1998; Kameswaran e Biegler, 2006). Esta falta de robustez deve-se ao mau condicionamento da linearização das equações do modelo (Ascher e Petzold, 1998), que resulta no mau condicionamento do problema de optimização (Santos, 2001), e a falhas na solução das equações do modelo resultantes da utilização de estimativas inadequadas. Estas falhas ocorrem porque as equações do modelo são resolvidas em ciclo aberto, ou seja, durante a solução das equações do modelo não são feitas quaisquer alterações à trajectória de controlo.

O problema do condicionamento pode ser atenuado pela utilização de compensadores internos que estabilizem o processo, ou atenuem a sua instabilidade (Santos et al., 1996; Gândara e Oliveira, 2001). As falhas na solução das equações do modelo podem ser ultrapassadas utilizando a técnica de *incremental single shooting* proposta por Michalik et al. (2009) para problemas de estimação de parâmetros. Estes autores observam que as falhas na solução das equações do modelo ocorrem, geralmente, nas iterações iniciais do problema, quando as estimativas dos parâmetros ainda podem estar bastante afastadas da solução, resultando em trajectórias inviáveis. Para ultrapassar este problema, é proposto que, no caso de se verificar uma falha, o problema seja primeiro resolvido num horizonte temporal mais pequeno. A solução para este problema, poderá ser suficiente para estabilizar o processo num horizonte mais longo, permitindo assim a solução incremental das equações do modelo.

De entre os métodos directos, os métodos sequenciais são aqueles que resultam em problemas de optimização de menor dimensão, o que permite a sua aplicação a problemas de grande dimensão. Hartwich e Marquardt (2010) descrevem a aplicação de um método sequencial a um problema com cerca de 12000 variáveis de estado, utilizando intervalos de amostragem adaptativos, de diferentes tamanhos, de forma a minimizar o esforço numérico requerido.

### **Multiple shooting**

O método de integração por tentativas, ou *multiple shooting* (Bock e Plitt, 1984; Santos et al., 1995; Leineweber, 1999; Schäfer et al., 2007), surgiu como forma de ultrapassar a falta de robustez apontada aos métodos sequenciais. Este é ainda um método sequencial, no sentido em que são primeiro resolvidas as equações do modelo e só depois é resolvido o problema de optimização. A principal diferença relativamente aos métodos anteriormente descritos (por vezes também designados por *single shooting*) está no facto de as equações do modelo não serem resolvidas para todo o horizonte

de forma contínua. Assim, nos métodos de *multiple shooting* o horizonte de previsão é dividido em  $N$  subintervalos, sendo as equações do modelo resolvidas de forma independente em cada um deles. A parametrização das variáveis de controlo é feita tal como nos métodos sequenciais, sendo que na discussão seguinte será considerada uma aproximação de ordem zero para estas variáveis.

A resolução das equações do modelo de forma independente em cada um dos  $N$  subintervalos torna necessária a introdução de novas variáveis

$$s^x = \begin{bmatrix} s_0^x & s_1^x & \cdots & s_N^x \end{bmatrix},$$

$$s^z = \begin{bmatrix} s_0^z & s_1^z & \cdots & s_N^z \end{bmatrix}$$

com  $s_i^x \in \mathbb{R}^{n_x}$  e  $s_i^z \in \mathbb{R}^{n_z}$ , que, no caso dos primeiros  $(N - 1)$  elementos, representam as condições iniciais em cada um dos intervalos, para as variáveis diferenciais e algébricas, respectivamente. Com estas condições iniciais, e para a trajectória de controlo  $\tilde{u}$ , as equações do modelo são resolvidas em cada um dos intervalos:

$$0 = F(\dot{x}(t), x(t), z(t), \tilde{u}_i, t), \quad t \in [t_i, t_{i+1}], \quad i = 0, \dots, N - 1$$

$$x(t_i) = s_i^x$$

$$z(t_i) = s_i^z$$

A solução destas equações são  $N$  trajectórias independentes,  $\tilde{x}_i(t, \tilde{u}_i, s_i)$  e  $\tilde{z}_i(t, \tilde{u}_i, s_i)$ . Tal como nos métodos sequenciais, também aqui é necessária a obtenção das derivadas de primeira ordem da solução das equações do modelo.

Para que a solução do problema seja consistente, é necessário garantir a continuidade das variáveis de estado diferenciais entre os diferentes intervalos considerados. Esta condição é assegurada pela utilização das restrições:

$$s_{i+1}^x = \tilde{x}_i(t_{i+1}, \tilde{u}_i, s_i), \quad i = 0, \dots, N - 1.$$

Estas restrições apenas necessitam de ser estritamente satisfeitas no final do problema de optimização, não sendo necessária a sua satisfação em todas as iterações do algoritmo de optimização. Este é um dos aspectos mais importantes dos métodos de *multiple shooting*, uma vez que permite lidar com estimativas inviáveis de  $u$  e  $s$  (Binder et al., 2001). Estas restrições têm um efeito semelhante à parametrização das variáveis de estado feita nos métodos simultâneos, sendo através delas trocada informação entre os algoritmos de controlo e de resolução das equações do modelo. São também estas restrições que fazem com que o *multiple shooting* seja considerado um híbrido entre os métodos sequenciais e simultâneos.

As restrições de trajectória podem ser transformadas de forma semelhante à considerada nos métodos sequenciais, ou seja, transformando-as num conjunto de restrições

pontuais. Utilizando os mesmos pontos considerados para a divisão do horizonte de previsão, obtemos:

$$0 \geq S_i(\tilde{u}, s_i) = S(s_i^x, s_i^z, \tilde{u}, t_i) \quad i = 0, \dots, N.$$

Novamente aqui, esta transformação apenas garante o cumprimento das restrições de trajectória nos  $N + 1$  pontos para os quais são escritas explicitamente, não permitindo assegurar o seu cumprimento em todo o horizonte.

Resumindo, a parametrização feita nos métodos de *multiple shooting* transforma o problema (2.4) no problema discreto:

$$\begin{aligned} \min_{\tilde{u}, s} \quad & \Psi(\tilde{u}, s) = \Theta(s_N^x, s_N^z, t_F) + \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} L(\tilde{x}(\tilde{u}, s_i, t), \tilde{z}(\tilde{u}, s_i, t), \tilde{u}, t) dt \\ \text{s.a} \quad & 0 = s_{i+1}^x - \tilde{x}_i(t_{i+1}, \tilde{u}_i, s_i), \quad i = 0, \dots, N - 1 \\ & 0 = F(\dot{x}(t), x(t), z(t), \tilde{u}_i, t), \quad t \in [t_i, t_{i+1}], \quad i = 0, \dots, N - 1 \\ & 0 = x(t_i) - s_i^x, \quad i = 0, \dots, N - 1 \\ & 0 = z(t_i) - s_i^z, \quad i = 0, \dots, N - 1 \\ & 0 \geq S(s_i^x, s_i^z, \tilde{u}, t_i), \quad i = 0, \dots, N \\ & 0 = M(s_0^x, s_0^z, t_0) \\ & 0 = N(s_N^x, s_N^z, t_F) \end{aligned}$$

A natureza sequencial deste métodos reside no facto de, em cada iteração da resolução do problema, serem resolvidas as equações do modelo utilizando as estimativas de  $u$  e  $s$  mais recentes. A solução destas equações, bem como as suas derivadas, são depois utilizadas na iteração seguinte do algoritmo de optimização.

Uma característica importante deste problema é a sua esparsidade e estrutura, que resulta do facto de as equações do modelo serem integradas de forma independente em cada um dos  $N$  intervalos considerados. A matriz Hessiana do problema tem uma estrutura diagonal por blocos, com blocos não nulos correspondentes apenas às variáveis locais  $s_i$  e  $\tilde{u}_i$  (Bock e Plitt, 1984). Esta característica permitiu o desenvolvimento de estratégias eficientes para a obtenção da solução do problema (Leineweber et al., 2003; Schäfer et al., 2007), nomeadamente através da utilização de métodos SQP parcialmente reduzidos.

### Métodos simultâneos

Nos métodos simultâneos todas as variáveis do problema são parametrizadas e os problemas de optimização e de solução das equações do modelo são resolvidos simultaneamente (Biegler, 2010). A parametrização é, muito frequentemente, efectuada

ada recorrendo a métodos de colocação ortogonal em elementos finitos (Tsang et al., 1975; Biegler, 1984; Logsdon e Biegler, 1989). São utilizados elementos finitos por, tal como no método sequencial, estes permitirem a utilização de aproximações polinomiais com grau mais baixo (Cuthrell e Biegler, 1987). O horizonte de previsão é dividido em  $N$  elementos finitos, de tamanho  $T$ , sendo considerados  $NC$  pontos de colocação no interior de cada elemento.

Em cada elemento finito é considerada uma aproximação polinomial das variáveis de estado diferenciais utilizando, por exemplo, polinómios de Lagrange

$$x_n^m(t) = \sum_{j=0}^{NC} \ell_j(t) \cdot x_{n,j}^m, \quad m = 1, \dots, n_x, \quad n = 1, \dots, NC$$

com:

$$\ell_j(t) = \prod_{\substack{i=0 \\ i \neq j}}^{NC} \frac{t - t_{n,i}}{t_{n,j} - t_{n,i}}$$

Aqui  $x_n^m(t)$  é a aproximação polinomial da variável de estado diferencial  $m$  no elemento  $n$ ,  $x_{n,j}^m$  é o valor dessa variável no ponto de colocação  $j$  do elemento  $n$  e  $t_{n,i}$  é o ponto de colocação  $i$  no elemento  $n$ . Utilizando esta representação, as variáveis de estado diferenciais e as suas derivadas em ordem ao tempo,  $\dot{x}$ , nos pontos de colocação são obtidas por

$$x_n^m(t_{n,i}) = \sum_{j=0}^{NC} \ell_j(t_{n,i}) \cdot x_{n,j}^m = x_{n,i}^m, \quad i = 1, \dots, NC$$

$$\frac{dx_n^m(t_{n,i})}{dt} = \sum_{j=0}^{NC} \frac{d\ell_j(t_{n,i})}{dt} \cdot x_{n,j}^m = \dot{x}_{n,i}^m, \quad i = 1, \dots, NC$$

A continuidade dos perfis das variáveis de estado diferenciais tem de ser assegurada entre os intervalos. Para isso, é conveniente a utilização dos pontos de colocação de Radau, em que o último ponto de colocação coincide com o instante final do elemento (Ascher e Petzold, 1998). Neste caso, as restrições de continuidade podem ser escritas como:

$$x_{n-1,NC} = x_{n,0}$$

Em alternativa, podem ser acrescentadas à formulação restrições que garantam a satisfação destas condições.

As variáveis de estado algébricas e de controlo podem ser discretizadas utilizando representações polinomiais diferentes das utilizadas para as variáveis diferenciais, mas os pontos de colocação devem ser os mesmos (Biegler, 2007). No caso destas variáveis, não são necessárias restrições de continuidade. Tal como nos restantes métodos, é habitual a utilização de aproximações de ordem zero para as variáveis de controlo.

As restrições de trajectória são discretizadas considerando o seu valor em todos os pontos de colocação no horizonte, podendo, para além destes, ser utilizado um conjunto adicional de pontos (Binder et al., 2001), sendo garantida a sua satisfação em todos os pontos considerados. Ainda que a satisfação destas restrições continue apenas a ser garantida num número finito de pontos, nos métodos simultâneos este número é bastante superior ao utilizado nos restantes métodos, devido à utilização de pontos de colocação no interior dos elementos finitos.

Nos métodos simultâneos as variáveis de decisão são o valor de todas as variáveis em todos os pontos de colocação de todos os elementos. Para simplificar a representação do problema serão utilizados os vectores:

$$\begin{aligned} X &= \begin{bmatrix} x_{1,1}^T & \cdots & x_{N,NC}^T \end{bmatrix}^T \in \mathbb{R}^{n_x \times NC \times N} \\ Z &= \begin{bmatrix} z_{1,1}^T & \cdots & z_{N,NC}^T \end{bmatrix}^T \in \mathbb{R}^{n_z \times NC \times N} \\ U &= \begin{bmatrix} u_1^T & \cdots & u_N^T \end{bmatrix}^T \in \mathbb{R}^{n_u \times NC \times N} \end{aligned}$$

A aplicação desta parametrização ao problema (2.4) resulta no problema:

$$\begin{aligned} \min_{X,Z,U} \quad & \Psi(X,Z,U) \\ \text{s.a} \quad & 0 = F(\dot{x}_{n,i}, x_{n,i}, z_{n,i}, u_n, t_{n,i}), \quad i = 1, \dots, NC, \quad n = 1, \dots, N \\ & 0 = x_{n-1,NC} = x_{n,0}, \quad n = 2, \dots, N \\ & 0 = M(x_{1,0}, z_{1,0}, t_0) \\ & 0 = N(x_{N,NC}, z_{N,NC}, t_0) \\ & 0 \geq S(x_{n,i}, z_{n,i}, u_n, t_{n,i}), \quad i = 1, \dots, NC, \quad n = 1, \dots, N \end{aligned}$$

Como se pode verificar, a utilização dos métodos simultâneos resulta em problemas de optimização de grande dimensão. Estes problemas podem ser resolvidos utilizando algoritmos especificamente desenhados para este fim, tirando partido da estrutura deste problema (Kameswaran e Biegler, 2006). Neste contexto, os métodos de ponto interior implementados na rotina IPOPT (Wächter, 2002; Wächter e Biegler, 2006) têm demonstrado particular eficiência. A utilização de métodos SQP parcialmente reduzidos (Cervantes et al., 2000; Cervantes e Biegler, 2000) é particularmente vantajosa quando o número de graus de liberdade do problema é relativamente pequeno. Já quando o número de graus de liberdade é elevado, será mais eficiente a utilização de métodos que, tirando partido da estrutura e esparsidade das matrizes, não decompõem o passo no algoritmo de optimização (Kameswaran e Biegler, 2006; Wächter, 2002).

Outro aspecto importante nos métodos simultâneos é a precisão com que as equações do modelo são resolvidas, sendo esta questão particularmente relevante para a aplicação a processos instáveis ou *stiff*. Nestes casos será necessária a divisão do



horizonte temporal numa grelha mais fina, o que resulta num problema de optimização de grande dimensão. Uma abordagem possível a este problema é o controlo do erro na solução das equações do modelo considerando intervalos finitos com dimensão variável. Neste caso, a dimensão dos intervalos é também uma variável de decisão do problema de optimização. Esta estratégia pode contudo resultar num aumento significativo da não-linearidade do problema correspondente (Biegler et al., 2002; Vasantharajan e Biegler, 1990). Outras estratégias para ajuste dos elementos finitos foram propostas, nomeadamente para a detecção de *switching conditions* (Biegler, 2007) e para a equidistribuição dos erros ao longo do horizonte (Vasantharajan e Biegler, 1990). Esta é também uma diferença importante entre estes métodos e os métodos sequenciais e de *multiple shooting*, em que este problema não se verifica. Nestes, a discretização ao nível do problema de optimização é independente da discretização utilizada para a integração das equações, bastando que ambas coincidam apenas nos extremos dos intervalos de discretização do problema de optimização. Assim, é possível a utilização de discretizações “grosseiras” ao nível do problema de optimização, sem que tal tenha quaisquer consequências na qualidade do resultado da integração das equações do modelo ou da sua informação de primeira ordem.

De entre as vantagens geralmente atribuídas aos métodos sequenciais, as mais importantes são a facilidade em lidar com restrições de trajectória em quaisquer das variáveis do problema, a capacidade de lidar com modos instáveis e o facto de não ser necessária a integração das equações do modelo em cada passo do algoritmo de solução (Biegler et al., 2002). A solução das equações do modelo apenas é conhecida depois de resolvido o problema de optimização, uma vez que as trajectórias correspondentes às iterações deste problema não satisfazem as equações do modelo.

### Métodos híbridos

Recentemente foram propostos métodos simultâneos híbridos, ou *quasi*-sequenciais, em que as restrições das equações do modelo são satisfeitas em todas as iterações (Hong et al., 2006; Bartl et al., 2011). Neste método as equações do modelo são resolvidas em cada iteração de optimização, por colocação em elementos finitos, tal como nos métodos simultâneos. As equações, do modelo e das suas sensibilidades, são resolvidas sucessivamente de elemento para elemento. A solução destas equações é depois utilizada para eliminar as variáveis de estado, resultando daqui que as variáveis de decisão do problema de optimização são apenas as variáveis de controlo. O problema de optimização é resolvida utilizando um método SQP reduzido, em que o passo no espaço nulo das equações do modelo corresponde à actualização das variáveis de controlo em cada iteração.

Relativamente aos métodos simultâneos, Hong et al. (2006) referem como principais vantagens deste método a maior facilidade na pesquisa linear, resultando na utiliza-



ção de passos completos com maior frequência, e a menor dimensão do problema de otimização que é resolvido, o que é conseguido pela eliminação das restrições de igualdade. Relativamente aos métodos sequenciais, as principais vantagens são a possibilidade de garantir a satisfação das restrições de trajectória num maior número de pontos, resultante do facto de as equações do modelo serem resolvidas por colocação. Este método é expandido por [Bartl et al. \(2011\)](#) de forma a melhorar a precisão na solução das equações do modelo. Esta melhoria é conseguida controlando o erro relativo através da adaptação dos tamanhos dos elementos finitos em que é feita colocação. Uma vez que a solução das equações do modelo é independente da solução do problema de otimização, a estratégia utilizada permite alterar a dimensão dos elementos, sem que seja necessário aumentar a dimensão do problema de otimização. Os resultados apresentados mostram que a melhoria da precisão na solução das equações do modelo permite uma redução significativa do número de iterações necessárias para a obtenção da solução.

## 2.4 Solução de DAE's e cálculo de sensibilidades

A resolução dos problema de otimização dinâmica por métodos sequenciais implica que, em cada iteração, seja obtida a solução das equações do modelo, bem como as derivadas de primeira ordem dessa solução. Nesta secção serão apresentados métodos numéricos para a obtenção desta informação. De notar que apenas serão discutidos métodos em que a solução das DAE's é obtida de forma progressiva, avançando no horizonte temporal, designados métodos *forward*.

Para esta discussão iremos considerar DAE's na forma implícita

$$F(\dot{v}, v, \theta, t) = 0 \quad (2.8)$$

com condições iniciais:

$$M(\dot{v}(t_0), v(t_0), \theta, t_0) = 0 \quad (2.9)$$

A solução numérica destas equações é sempre obtida discretizando a variável tempo, ou seja, a solução obtida não é contínua, sendo apenas calculada num conjunto finito de instantes de tempo. Nos métodos em que esta solução é obtida de forma progressiva, a solução para o instante  $t_{(n+1)}$  é calculada depois de obtida a solução para o instante  $t_{(n)}$ , sendo a diferença entre os dois instantes de tempo designada passo de integração. Este passo de integração pode variar ao longo do horizonte temporal da integração para, por exemplo, controlar o erro local ([Ascher e Petzold, 1998](#)).

Para que a solução de (2.8) seja possível, é necessária a sua transformação num problema algébrico, o que pode ser feito substituindo, em cada passo de integração, os termos diferenciais,  $\dot{v}$ , por uma aproximação baseada em diferenças. Na obtenção da

solução em  $t_{(n+1)}$  esta aproximação pode ser feita utilizando apenas a solução obtida para  $t_{(n)}$ , ou fazendo interpolação com as soluções obtidas nos instantes anteriores. Os métodos que utilizam esta última aproximação são genericamente designados métodos multipasso, e entre eles estão os métodos que utilizam fórmulas de diferenças anteriores (*backward difference formula*, BDF). Utilizando uma BDF de ordem  $k$  e passo  $h$  no instante  $t_{(n)}$ , obtém-se o sistema de equações algébricas não lineares (Ascher e Petzold, 1998)

$$g(v_{(n+1)}) = F \left( \sum_{j=0}^k \alpha_j v_{(n+1-j)}, v_{(n+1)}, \theta, t_{(n+1)} \right) = 0$$

onde  $\alpha_j$  e  $k$  são coeficientes escalares da fórmula. A solução  $v_{(n+1)}$  pode ser obtida iterativamente pelo método de Newton. Na iteração  $m$  o sistema linear que precisa de ser resolvido é

$$J \left( v_{(n+1)}^{(m)} - v_{(n+1)}^{(m+1)} \right) = g \left( v_{(n+1)}^{(m)} \right) \quad (2.10)$$

onde:

$$J = \alpha_0 \frac{\partial F}{\partial \dot{v}_{(n+1)}^{(m)}} + \frac{\partial F}{\partial v_{(n+1)}^{(m)}} \quad (2.11)$$

O cálculo do Jacobiano,  $J$ , em cada iteração do método de Newton pode ser uma tarefa computacionalmente dispendiosa, que pode ser evitada utilizando um Jacobiano  $\hat{J}$  calculado numa iteração anterior, até que a velocidade de convergência não seja satisfatória (Brenan et al., 1989). Quando tal ocorre, o Jacobiano é novamente calculado. A implementação mais popular deste classe de métodos é a rotina DASSL (Petzold, 1982). Uma implementação mais recente destes métodos está disponível na biblioteca SUNDIALS (Hindmarsh et al., 2005).

Para o sistema (2.8) podemos definir  $n_s = n_v \cdot n_p$  sensibilidades das variáveis de estado relativamente aos parâmetros

$$s_i = \frac{\partial v}{\partial \theta_i}, \quad i = 1, \dots, n_p$$

onde  $s_i \in \mathbb{R}^{n_v}$  é o vector das sensibilidades das variáveis de estado relativamente ao parâmetro  $\theta_i$ . Por diferenciação parcial de (2.8) e (2.9), obtemos

$$\frac{\partial F}{\partial v} s_i + \frac{\partial F}{\partial \dot{v}} \dot{s}_i + \frac{\partial F}{\partial \theta_i} = 0, \quad i = 1, \dots, n_p$$

com condições iniciais:

$$\frac{\partial M}{\partial v} s_i(t_0) + \frac{\partial M}{\partial \dot{v}} \dot{s}_i(t_0) + \frac{\partial M}{\partial \theta_i} = 0, \quad i = 1, \dots, n_p$$

que constituem as equações das sensibilidades. Estas são sempre lineares, indepen-

dentemente da não linearidade das DAE's a partir das quais são obtidas (Brenan et al., 1989).

De entre os métodos numéricos existentes para a solução das equações das sensibilidades serão abordados os métodos alternado directo (Caracotsios e Stewart, 1985), o alternado corrector<sup>1</sup> (Feehery et al., 1997) e o simultâneo corrector (Maly e Petzold, 1996), baseados na solução numérica das DAE's recorrendo a BDF's. As principais diferenças entre estes métodos residem na forma como o Jacobiano do problema (2.8) é aproveitado para a resolução das equações das sensibilidades e quando estas são resolvidas. Todos estes métodos estão implementados na rotina DASPK (Li e Petzold, 1999), uma extensão da já referida rotina DASSL.

No método alternado directo, as sensibilidades são calculadas apenas após a convergência da solução das DAE's. As sensibilidades são calculadas através da solução directa de

$$A \begin{bmatrix} \dot{s}_i \\ s_i \end{bmatrix} + \frac{\partial F}{\partial \theta_i} = 0 \quad (2.12)$$

onde:

$$A = \begin{bmatrix} \frac{\partial F}{\partial \dot{v}_{(n+1)}}, & \frac{\partial F}{\partial v_{(n+1)}} \end{bmatrix}.$$

As variáveis  $\dot{s}_i$  são eliminadas por utilização de uma BDF. Este método requer, em cada passo da integração, a obtenção e factorização da matriz  $A$ , apesar de esta ser construída a partir da mesma informação utilizada para construir a matriz  $\hat{J}$ . A razão para tal é o facto de esta última apenas ser actualizada ocasionalmente.

Tal como no método anterior, no método alternado corrector as equações das sensibilidades apenas são resolvidas após convergência da solução das DAE's. De forma a evitar a factorização da matriz  $A$  em cada passo do integrador, no método alternado corrector as equações (2.12) são resolvidas iterativamente utilizando um método *quasi*-Newton:

$$\hat{J} \begin{pmatrix} s_{i(n+1)}^{(m)} - s_{i(n+1)}^{(m+1)} \end{pmatrix} = A \begin{bmatrix} \alpha_0 s_{i(n+1)}^{(m)} + \sum_{j=1}^k \alpha_j s_{i(n+1-j)}^{(m)} \\ s_{i(n+1)}^{(m)} \end{bmatrix} + \frac{\partial F}{\partial \theta_i} \quad (2.13)$$

Apesar de continuar a ser necessário, em cada passo do integrador, a obtenção da matriz  $A$ , neste método a sua factorização não é necessária. A factorização necessária é a da matriz  $\hat{J}$ , já obtida para a resolução de (2.10).

Ao contrário dos métodos anteriores, no método simultâneo corrector, as DAE's e as equações das sensibilidades são resolvidas simultaneamente. Para tal, definem-se os

---

<sup>1</sup>Staggered direct e staggered corrector, na designação original

vectores alargados:

$$\mathcal{F} = \left[ F(\dot{v}, v, \theta, t), \frac{\partial F}{\partial v} s_1 + \frac{\partial F}{\partial \dot{v}} \dot{s}_1 + \frac{\partial F}{\partial \theta_1}, \dots, \frac{\partial F}{\partial x} s_{n_p} + \frac{\partial F}{\partial \dot{v}} \dot{s}_{n_p} + \frac{\partial F}{\partial \theta_{n_p}} \right]^T$$

$$X = [v, s_1, \dots, s_{n_p}]$$

Utilizando estes vectores, as DAE's e as equações das sensibilidades podem ser escritas como:

$$\mathcal{F}(\dot{X}, X, \theta, t) = 0, \quad \Phi(\dot{X}(t_0), X(t_0), \theta, t_0) = 0$$

Utilizando uma BDF de ordem  $k$  para eliminação dos termos diferenciais, obtém-se um sistema de  $n_v \cdot (n_p + 1)$  equações algébricas:

$$\mathcal{G}(X_{(n+1)}) = \mathcal{F} \left( \sum_{j=0}^k \alpha_j X_{(n+1-j)}, X_{(n+1)}, \theta, t \right) = 0$$

Para a solução deste sistema pelo método de Newton é necessário, em cada iteração  $m$  deste método, resolver o sistema de equações lineares

$$M \left( X_{(n+1)}^{(m)} - X_{(n+1)}^{(m+1)} \right) = \mathcal{G} \left( X_{(n+1)}^{(m)} \right)$$

onde

$$M = \begin{bmatrix} J & & & & & \\ J_1 & J & & & & \\ J_2 & 0 & J & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ J_{n_p} & 0 & \cdots & 0 & J & \end{bmatrix}$$

com  $J$  definido por (2.11) e:

$$J_i = \frac{\partial J}{\partial v} s_i + \frac{\partial J}{\partial \theta_i}, \quad i = 1, \dots, n_p$$

A aplicação directa deste método resulta na necessidade, em cada iteração do método de Newton, do cálculo das matrizes  $J_i$ . No entanto, [Maly e Petzold \(1996\)](#) mostram que se  $M$  for aproximada pela sua parte diagonal por blocos, a ordem de convergência do método de Newton é *2-step quadratic*. Tal permite a utilização da aproximação  $\hat{J}$  e a reutilização de  $M$  em diferentes passos da integração. Este método elimina a necessidade da factorização de matrizes adicionais para o cálculo das sensibilidades, mas requer o cálculo dos resíduos de (2.4) em cada iteração do método de Newton.

Os resíduos das sensibilidades são a diferença entre o seu valor calculado por qualquer um dos métodos e o seu valor real. Recorrendo a (2.12) verificamos que estes resíduos

são

$$r_i = A \begin{bmatrix} \dot{s}_i \\ s_i \end{bmatrix} + \frac{\partial F}{\partial \theta_i}, \quad i = 1, \dots, n_p$$

Se for possível calcular as derivadas  $\partial F / \partial \theta_i$  analiticamente, então a avaliação destes resíduos é directa. Se tal não for possível existem ainda duas alternativas. Uma delas é a aproximação das derivadas por diferenças finitas, sendo neste caso os resíduos calculados por

$$r_i = A \begin{bmatrix} \dot{s}_i \\ s_i \end{bmatrix} + \frac{F(\dot{v}, v, \theta + \delta_i e_i, t) - F(\dot{v}, v, \theta, t)}{\delta_i}$$

onde  $\delta_i$  é uma quantidade escalar pequena e  $e_i$  é o vector unidade  $i$ . [Maly e Petzold \(1996\)](#) mostram que o cálculo dos resíduos das sensibilidades também é possível utilizando uma aproximação por diferenças finitas de derivadas direccionais:

$$r_i = \frac{F(\dot{v} + \delta_i \dot{s}_i, v + \delta_i s_i, \theta + \delta_i e_i, t) - F(\dot{v}, v, \theta, t)}{\delta_i}$$

Esta aproximação tem a vantagem de não ser necessário obter a matriz  $A$  para o cálculo dos resíduos, logo, o esforço computacional necessário é inferior ao necessário na utilização de diferenças finitas. Este método tem no entanto a desvantagem de ser difícil a análise da sua precisão, sendo crucial a escolha de  $\delta_i$ . O cálculo dos resíduos desta forma apenas é aplicável aos métodos simultâneo corrector e alternado corrector, uma vez que nestes não é necessário formar explicitamente o sistema de equações lineares (2.12).

Para além das técnicas apresentadas, baseadas em BDF's, foram propostos outros métodos para a obtenção da solução numérica de DAE's, entre os quais os baseados na discretização de Euler. Um destes métodos foi proposto por [Deuffhard et al. \(1987\)](#) para DAE's na forma linear implícita, em que, conjuntamente com a discretização de Euler, são utilizadas técnicas de extrapolação. Utilizando uma estratégia semelhante ao método simultâneo corrector, este método foi expandido por [Schlegel et al. \(2004\)](#) para a solução simultânea das equações do modelo e das sensibilidades, sendo adequado à solução de problemas de muito grande dimensão ([Hartwich e Marquardt, 2010](#)). Uma implementação paralelizável deste método é descrita por [Hartwich et al. \(2010\)](#).

Os métodos *forward* descritos são mais adequados quando se pretende obter a sensibilidade de um grande número de variáveis relativamente a um conjunto pequeno de parâmetros. Na situação inversa, ou seja, para obtenção da sensibilidade de um conjunto pequeno de variáveis relativamente a um grande número de parâmetros, poderá ser mais eficiente a utilização do método das equações adjuntas ([Cao et al., 2003](#)). Neste método, as equações do modelo são primeiro resolvidas, partindo do instante inicial até ao instante final. As equações adjuntas, a partir das quais são obtidas as sensibilidades, são depois resolvidas de forma regressiva, ao longo da trajectória ob-

tida na solução das equações do modelo. A existência de restrições de trajectória para o estados torna mais complexa a aplicação deste método. É também necessário o armazenamento de grandes quantidades de informação, o que torna este método pouco eficiente para aplicação a problemas de grande dimensão (Hartwich et al., 2010).

## 2.5 Solução do problema de optimização

Todas as estratégias de resolução de problemas de optimização dinâmica por métodos directos transformam o problema de optimização inicial num problema (ou sequência de problemas) de programação não-linear (NLP) com restrições. De forma genérica, este problema pode ser escrito como

$$\min_x f(x) \quad (2.14a)$$

$$\text{s.a } c_i(x) = 0, \quad i \in \mathcal{E} \quad (2.14b)$$

$$c_i(x) \geq 0, \quad i \in \mathcal{I} \quad (2.14c)$$

onde  $x \in \mathbb{R}^n$  é o vector das variáveis de decisão,  $f$  é a função objectivo escalar,  $c_i$ ,  $i \in \mathcal{E}$ , é o conjunto das restrições de igualdade e  $c_i$ ,  $i \in \mathcal{I}$ , é o conjunto das restrições de desigualdade.

### 2.5.1 Programação quadrática sucessiva

A classe de métodos mais popular para a resolução de problemas de programação não-linear em engenharia de processos é a programação quadrática sucessiva (*sequential quadratic programming*, SQP). As principais razões para a sua vasta aplicação são o facto de se ter demonstrado que necessitam de menos avaliações das funções e poderem ser adaptados à estrutura de problemas específicos (Biegler e Grossmann, 2004).

Os métodos SQP obtêm uma solução para o problema (2.14) resolvendo, em cada iteração  $k$ , um problema quadrático em torno de  $x_k$ :

$$d_k = \arg \min_d \frac{1}{2} d^T W_k d + \nabla f_k^T d \quad (2.15a)$$

$$\text{s.a } A_{\mathcal{E}}(x_k) d + c_{\mathcal{E}}(x_k) = 0 \quad (2.15b)$$

$$A_{\mathcal{I}}(x_k) d + c_{\mathcal{I}}(x_k) \geq 0 \quad (2.15c)$$

com  $d \in \mathbb{R}^n$ , e onde  $A_{\mathcal{E}}$  e  $A_{\mathcal{I}}$  são, respectivamente, os Jacobianos das restrições de

igualdade e desigualdade:

$$\begin{aligned} A_{\mathcal{E}}^{(i)} &= \nabla c_i^T, & i \in \mathcal{E} \\ A_{\mathcal{I}}^{(i)} &= \nabla c_i^T, & i \in \mathcal{I} \end{aligned}$$

A matriz  $W$  é a Hessiana da função Lagrangiana,  $\nabla_{xx}^2 \mathcal{L}(x, \lambda)$ , definida como

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T c(x)$$

sendo  $\lambda$  o vector dos multiplicadores de Lagrange das restrições. Se  $d_k$  e  $\lambda_k$  forem a solução de (2.15), a nova iteração do problema (2.14) é definida por  $x_{k+1} = x_k + d_k$  e  $\lambda_{k+1} = \lambda_k$ .

Se  $x^*$  for a solução de (2.14) então, sob certas condições e se  $(x_k, \lambda_k)$  for suficientemente próximo de  $(x^*, \lambda^*)$ , o conjunto de restrições activas,  $\mathcal{A}_k$ , do problema quadrático (2.15) é igual ao conjunto de restrições activas,  $\mathcal{A}(x^*)$ , do problema não-linear (2.14) na solução (Nocedal e Wright, 1999). Nestas condições a solução do problema quadrático é equivalente a uma iteração do método de Newton aplicado às condições de optimalidade de Karush-Kuhn-Tucker (KKT) do problema não-linear. É a esta equivalência que se devem as propriedades de convergência local (bastante rápida) dos métodos SQP.

A resolução do problema quadrático, no âmbito dos métodos SQP é discutida na secção seguinte. Após esta etapa, de forma a assegurar a convergência dos métodos SQP a partir de soluções iniciais afastadas da solução óptima, é necessário controlar o tamanho do passo  $d_k$ , o que pode ser feito recorrendo a métodos de globalização.

### Resolução do problema quadrático

A solução do problema quadrático pode ser obtida identificando, iterativamente, o conjunto de restrições activas na solução do QP, tratando-as depois como igualdades. Este é um problema combinatório, cuja complexidade aumenta rapidamente com o número de restrições. Este problema pode ser atenuado recorrendo a procedimentos de *hot start* (ou *warm start*), em que a estimativa inicial do passo é igual à solução do último QP resolvido e em que a estimativa inicial do conjunto de restrições activas é igual ao do NLP na iteração anterior. Estes procedimentos são particularmente eficazes quando aplicados em problemas de programação dinâmica com horizontes regressivos (como, por exemplo, controlo predictivo), em que as diferenças entre dois problemas consecutivos poderão ser relativamente pequenas. Também como já foi referido, na vizinhança da solução do problema não linear, o conjunto de restrições activas do QP é igual ao do problema não linear. Em consequência, o conjunto de restrições activas não se altera de uma iteração para a outra do método SQP, sendo por isso conhecido, à partida, o conjunto de restrições activas do QP, podendo este ser

facilmente resolvido.

O problema da identificação do conjunto correcto de restrições activas pode ser evitado pela utilização de métodos de barreira como, por exemplo, os métodos de ponto interior discutidos na secção 2.5.5. A utilização de informação relativamente ao conjunto de restrições activas na solução em métodos de ponto interior é uma questão ainda em desenvolvimento (Gondzio e Grothey, 2008; Benson e Shanno, 2008).

A solução do problema quadrático pode ser obtida resolvendo directamente o problema (2.15) e obtendo simultaneamente todos os elementos do passo. Estes métodos são usualmente designados *full-space*. Em alternativa, a solução do problema quadrático pode ser obtida dividindo o passo em componentes mais pequenos, tal como nos métodos *reduced-space* (Biegler, 2010).

Neste caso, a solução do problema (2.15) pode ser dividida em duas componentes:

$$d_k = Y_k d_Y + Z_k d_Z \quad (2.16)$$

onde  $Y_k d_Y$  é a componente do passo no espaço característico e  $Z_k d_Z$  é a componente no espaço nulo das restrições de igualdade do problema. A matriz  $Z_k$  deve ser tal que as suas colunas constituam uma base do espaço nulo das restrições de igualdade do problema (2.15), ou seja:

$$A_g(x_k)Z_k = 0 \quad (2.17)$$

Substituindo em (2.15b), obtém-se

$$A_g(x_k)Y_k d_Y = -c_g(x_k) \quad (2.18)$$

ou seja, a componente do passo no espaço característico é:

$$d_Y = -[A_g(x_k)Y_k]^{-1}c_g(x_k)$$

Substituindo (2.16) no problema quadrático (2.15), e considerando (2.17), verifica-se que a componente no espaço nulo pode ser obtida por resolução do problema quadrático

$$\min_{d_Z} \left( Z_k^T \nabla f_k + Z_k^T W_k Y_k d_Y \right)^T d_Z + \frac{1}{2} d_Z^T Z_k^T W_k Z_k d_Z \quad (2.19a)$$

$$\text{s.a } A_g Z_k d_Z \geq -c_g(x_k) - A_g Y_k d_Y \quad (2.19b)$$

onde foram omitidos os termos constantes. O número de graus de liberdade deste problema quadrático é inferior ao do problema original, mantendo-se o número de restrições de desigualdade e, portanto, o problema de identificação do conjunto de restrições activas. As restrições de desigualdade deste problema assumem uma forma bastante mais simples quando, no problema original, as restrições de trajectória são



apenas limites aos valores que as variáveis podem tomar (Wächter, 2002).

Na ausência de restrições, a solução de (2.19) é obtida pela solução do sistema linear

$$(Z_k^T W_k Z_k) d_Z = - \left( Z_k^T \nabla f_k + Z_k^T W_k Y_k d_Y \right)$$

quando a matriz  $Z_k^T W_k Z_k$  é positiva definida. Uma forma comum de aproximar o termo cruzado  $Z_k^T W_k Y_k d_Y$  é considerá-lo nulo. Esta aproximação é tanto melhor quanto menor for a componente do passo no espaço característico, sendo particularmente útil para problemas com restrições lineares para as quais  $x_k$  é viável. Nesta situação  $d_Y$  é nulo, anulando-se também o termo cruzado. Esta aproximação também é adequada quando  $Z_k$  e  $Y_k$  são representações ortogonais, uma vez que  $Y_k d_Y$  é minimizado. Noutras condições, a contribuição do espaço característico pode ser importante, resultando a aproximação do termo cruzado por zero em convergência errática (Biegler, 2010). Nesta situação podem ser utilizadas diversas outras aproximações a este termo (Biegler et al., 1995).

A escolha das matrizes  $Y$  e  $Z$  pode ser feita de várias formas. Uma das possibilidades é a utilização de uma factorização ortogonal. Considerando  $A_g \in \mathbb{R}^{m \times n}$ , com  $m < n$ , não-singular, a factorização QR da sua transposta resulta em (Golub e Van Loan, 1996)

$$A_g^T(x_k) \Pi = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad (2.20)$$

onde  $Q \in \mathbb{R}^{n \times n}$  é uma matriz ortogonal,  $R_1 \in \mathbb{R}^{m \times m}$  é uma matriz triangular superior e  $\Pi \in \mathbb{R}^{m \times m}$  é uma matriz de permutações. As colunas das submatrizes  $Q_1 \in \mathbb{R}^{n \times m}$  e  $Q_2 \in \mathbb{R}^{n \times (n-m)}$  são ortonormais. A matriz de permutações é construída de forma a que  $R_1$  seja, não-singular e, tanto quanto possível, bem condicionada. A partir desta factorização, podemos definir as matrizes  $Y$  e  $Z$  como

$$\begin{bmatrix} Y_k & Z_k \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$$

Considerando (2.20) e substituindo em (2.18) verificamos que a componente  $d_Y$  do passo é solução de

$$R_1^T d_Y = -\Pi^T c_g(x_k)$$

obtida por uma simples substituição progressiva. A principal desvantagem desta escolha é o custo computacional da factorização QR, mesmo em sistemas esparsos. A principal vantagem é a sua robustez, uma vez que, como já referido, a utilização da factorização QR com permutações garante que  $R_1$  é, pelo menos, tão bem condicionada como  $A_g$ .

Alternativamente, as matrizes  $Y$  e  $Z$  podem ser obtidas por divisão das variáveis do problema em dependentes e independentes e aplicando um método de eliminação generalizado (Fletcher, 1987). Considerando  $A_g \in \mathbb{R}^{m \times n}$ , as variáveis do problema

podem ser divididas em  $x_D \in \mathbb{R}^m$  variáveis dependentes e  $x_I \in \mathbb{R}^{n-m}$  variáveis independentes

$$x = \begin{bmatrix} x_D^T & x_I^T \end{bmatrix}^T$$

o que corresponde a uma partição das colunas do Jacobiano como:

$$A_g(x_k) = \begin{bmatrix} C(x_k) & N(x_k) \end{bmatrix}$$

Com esta partição, as matrizes base pode ser definidas como (Fletcher, 1987)

$$Y_k = \begin{bmatrix} I \\ 0 \end{bmatrix}; \quad Z_k = \begin{bmatrix} -C_k^{-1}N_k \\ I \end{bmatrix}$$

onde  $C_k = C(x_k)$  e  $N_k = N(x_k)$ . A matriz  $C_k$  deve ser não singular e bem condicionada. Esta escolha das matrizes base tem algumas vantagens, entre elas o facto do sistema (2.18) se reduzir a

$$C_k d_Y = -c_g(x_k) \quad (2.21)$$

o que permite explorar a esparsidade e estrutura da matriz  $C_k$ , frequentes em problemas de optimização dinâmica de processos químicos. Outra vantagem é a possibilidade de obter as  $m \times (n-m)$  linhas superiores de  $Z_k$  utilizando as colunas de  $N_k$  como lado direito de (2.21) e, portanto, utilizando a mesma factorização necessária para o cálculo de  $d_Y$ .

Se, no algoritmo de optimização, forem necessárias estimativas dos multiplicadores das restrições de igualdade do problema, estas podem ser obtidas directamente da solução do QP, no caso dos métodos *full-space*. Quando a solução do QP é obtida por decomposição do passo, as estimativas destes multiplicadores podem ser obtidas por (Biegler et al., 1995)

$$\lambda_k = - \left[ Y_k^T A_g^T(x_k) \right]^{-1} Y_k^T \nabla f_k$$

Uma descrição detalhada da implementação deste tipo de métodos pode ser encontrada, por exemplo, em Biegler et al. (1995), e nas referências aí citadas. Para além da redução da dimensão do problema de optimização, esta metodologia tem diversas vantagens, uma vez que a estrutura da solução do problema de optimização fica evidenciada (Oliveira e Biegler, 1994b).

### Condicionamento numérico

Qualquer que seja o método utilizado, a solução dos problemas de optimização é sempre obtida através da solução de sistemas de equações lineares. É por isso importante avaliar a robustez dos métodos utilizados na solução destes sistemas, ou seja, como é que as soluções produzidas variam na presença de perturbações que podem ser, por exemplo, devidas à utilização de aproximações ou a erros numéricos. Esta avaliação

pode ser feita através da avaliação do condicionamento numérico das matrizes dos coeficientes destes sistemas lineares, que pode ser quantificado pelo seu número de condição. Este aspecto é particularmente relevante no contexto da otimização dinâmica, uma vez que as matrizes resultantes da integração de modelos constituídos por DAE's podem ser extremamente mal condicionadas, mesmo em sistemas estáveis (Ascher e Petzold, 1998). Este problema é agravado em sistemas instáveis ou *stiff* (quando o modelo inclui termos com magnitudes de escalas de tempo muito diferentes).

Considerando um sistema genérico de equações lineares,  $Ax = b$ , com  $A \in \mathbb{R}^{n \times n}$  e  $x, b \in \mathbb{R}^n$ , a sensibilidade da solução pode ser avaliada recorrendo ao sistema parametrizado

$$(A + \epsilon F)x(\epsilon) = b + \epsilon f \quad (2.22)$$

com  $F \in \mathbb{R}^{n \times n}$ ,  $b, f \in \mathbb{R}^n$ , e  $\epsilon \in \mathbb{R}$ . O erro relativo da solução deste sistema é (Golub e Van Loan, 1996)

$$\frac{\|x(\epsilon) - x(0)\|}{\|x(0)\|} = \|A\| \|A^{-1}\| (\rho_A + \rho_b) + O(\epsilon^2) \quad (2.23)$$

em que  $\rho_A$  e  $\rho_b$  representam os erros relativos em  $A$  e  $b$ , respectivamente:

$$\rho_A = |\epsilon| \frac{\|F\|}{\|A\|}, \quad \rho_b = |\epsilon| \frac{\|f\|}{\|b\|}$$

O número de condição da matriz,  $\kappa(A)$ , é definido como

$$\kappa(A) = \|A\| \|A^{-1}\|$$

correspondendo  $\kappa(A) = \infty$  a matrizes singulares. O valor do número de condição obtido depende da norma utilizada. No caso da norma euclidiana, o número de condição pode ser calculado através de (Golub e Van Loan, 1996)

$$\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$$

em que  $\sigma_{\max}(A)$  e  $\sigma_{\min}(A)$  são, respectivamente, os valores singulares máximo e mínimo da matriz  $A$ . Será esta a definição utilizada neste trabalho, pelo que o subscrito 2 deixará de ser apresentado.

Esta definição também é válida para matrizes rectangulares genéricas  $A \in \mathbb{R}^{n \times m}$ , quando  $\text{rank}(A) = m$ . Para qualquer matriz  $A$ ,  $\kappa(A) \geq 1$ , sendo esta matriz perfeitamente condicionada quando  $\kappa(A) = 1$ , ou seja, quando  $A$  é ortogonal. O condicionamento numérico do problema (2.22) é tanto pior quanto maior for o número de condição da matriz correspondente. Através de (2.23), facilmente se verifica que

quanto maior for o número de condição, mais sensível será a solução do sistema linear a perturbações. O número de condição de uma matriz também pode ser interpretado como uma medida da independência das suas colunas ou linhas. Se as colunas, ou linhas, de  $A$  forem quase linearmente dependentes, o seu número de condição será sempre elevado.

## 2.5.2 Aproximações da Hessiana

A matriz Hessiana  $W_k$  do problema (2.15) é definida como a Hessiana da função Lagrangiana, sendo esta definição que leva à equivalência dos métodos SQP com o método de Newton. No entanto, a informação de segunda ordem necessária para a sua construção pode não estar disponível, ou a sua obtenção ser dispendiosa do ponto de vista do esforço de cálculo necessário. Quanto tal ocorre, é necessário recorrer a aproximações, algumas das quais são discutidas de seguida. A utilização destas aproximações, no âmbito da otimização sem restrições, levou ao surgimento dos métodos designados *quasi-Newton*. Com algumas modificações, as mesmas aproximações podem ser utilizadas no âmbito da programação quadrática sucessiva.

### Método BFGS

Na resolução de problemas de otimização sem restrições, a aproximação mais comum é a denominada BFGS (Broyden, Fletcher, Goldfarb e Shanno), desenvolvida a partir de uma aproximação inicialmente proposta por Davidon (Nocedal e Wright, 1999). A ideia básica deste método é criar uma aproximação que utilize informação sobre a curvatura do problema, obtida a partir das derivadas de primeira ordem calculadas nas iterações mais recentes do algoritmo de otimização. Em problemas de otimização sem restrições, a utilização do método BFGS para aproximação da Hessiana permite uma convergência superlinear.

Definindo as quantidades

$$d_k = x_{k+1} - x_k, \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

pretende-se obter uma aproximação da Hessiana,  $B_{k+1}$ , simétrica e positiva definida que satisfaça a condição:

$$B_{k+1}d_k = y_k \tag{2.24}$$

Sendo  $B_{k+1}$  positiva definida, esta equação só terá solução quando for satisfeita a condição de curvatura:

$$d_k^T y_k > 0 \tag{2.25}$$

Pela imposição de algumas condições adicionais, pode obter-se uma solução única

para (2.24), dada por

$$B_{k+1} = (I - \rho_k y_k d_k^T) B_k (I - \rho_k d_k y_k^T) + \rho_k y_k y_k^T \quad (2.26)$$

onde  $\rho_k$  é um escalar obtido por:

$$\rho_k = \frac{1}{y_k^T d_k}$$

A equação (2.26) é conhecida como fórmula DFP (Davidon, Fletcher e Powell).

Os mesmos argumentos que levaram ao surgimento da fórmula DFP podem ser aplicados directamente à inversa da Hessiana  $H$  resultando na fórmula BFGS:

$$H_{k+1} = (I - \rho_k d_k y_k^T) H_k (I - \rho_k y_k d_k^T) + \rho_k d_k d_k^T. \quad (2.27)$$

Partindo de uma aproximação inicial da Hessiana,  $B_0$ , ou da sua inversa,  $H_0$ , simétricas e positivas definidas, as fórmulas DFP e BFGS geram sequências de matrizes também simétricas e positivas definidas.

No âmbito dos métodos SQP, a aplicação das fórmulas apresentadas deve ser adaptada, devido ao facto de a matriz que se pretende aproximar ser a Hessiana da função Lagrangiana,  $\nabla_{xx}^2 \mathcal{L}(x, \lambda)$ . Neste casos,  $d_k$  e  $y_k$  podem ser definidos como:

$$d_k = x_{k+1} - x_k, \quad y_k = \nabla_x \mathcal{L}(x_{k+1}, \lambda_{k+1}) - \nabla_x \mathcal{L}(x_k, \lambda_{k+1}). \quad (2.28)$$

No entanto, com esta definição não é possível garantir que, mesmo na vizinhança da solução, a condição de curvatura (2.25) é satisfeita. Logo, não é possível a aplicação directa da fórmula BFGS. É por isso necessária a utilização de outras estratégias, como o método BFGS amortecido. Definindo

$$r_k = \vartheta_k y_k + (1 - \vartheta_k) B_k d_k$$

com  $y_k$  e  $d_k$  definidos por (2.28) e o escalar  $\vartheta_k$  dado por:

$$\vartheta_k = \begin{cases} 1 & \text{se } d_k^T y_k \geq 0, 2d_k^T B_k d_k \\ (0, 8d_k^T B_k d_k) / (d_k^T B_k d_k - d_k^T y_k) & \text{se } d_k^T y_k < 0, 2d_k^T B_k d_k \end{cases}$$

A aproximação à matriz Hessiana é actualizada pela fórmula

$$B_{k+1} = B_k + \frac{B_k d_k d_k^T B_k}{d_k^T B_k d_k} + \frac{r_k r_k^T}{d_k^T r_k},$$

semelhante à fórmula BFGS para actualização da Hessiana, com  $y_k$  substituído por  $r_k$ . A designação BFGS amortecido deve-se ao facto de a aproximação à Hessiana obtida

por este método ser uma interpolação entre a aproximação obtida pela fórmula BFGS (quando  $\vartheta_k = 1$ ) e a manutenção da aproximação actual ( $B_{k+1} = B_k$ , quando  $\vartheta_k = 0$ ). A escolha de  $\vartheta_k$  é feita de forma a garantir que a nova aproximação obtida se mantém suficiente próxima de  $B_k$ , de forma a garantir que é positiva definida.

### Hessiana reduzida

Nos métodos em que a solução do QP é obtida por decomposição do passo é conveniente obter uma aproximação Hessiana do problema (2.19a),  $Z_k^T \nabla_{xx}^2 \mathcal{L}(x, \lambda) Z_k$ , correspondente à Hessiana projectada no espaço nulo das restrições de igualdade do problema. A aproximação à Hessiana projectada pode também ser obtida por utilização da fórmula BFGS, tendo sido propostas, para esse efeito, diferentes definições dos vectores  $s_k$  e  $y_k$  (Nocedal e Wright, 1999). Uma definição possível é:

$$\begin{aligned} d_k &= \alpha_k d_Z \\ y_k &= Z_k^T [\nabla_x \mathcal{L}(x_k + \alpha_k d_k, \lambda_{k+1}) - \nabla_x \mathcal{L}(x_k, \lambda_{k+1})] \end{aligned}$$

Uma vez que o que se pretende é aproximar a projecção da Hessiana no espaço nulo das restrições, foi proposto utilizar apenas a informação de curvatura obtida ao longo da componente  $d_Z$  do passo. Tal resulta em (Nocedal e Wright, 1999):

$$y_k = Z_k^T [\nabla_x \mathcal{L}(x_k + Z_k d_Z, \lambda_{k+1}) - \nabla_x \mathcal{L}(x_k, \lambda_{k+1})]$$

Esta última definição tem a desvantagem de ser necessária a avaliação dos gradientes da função objectivo e das restrições no ponto intermédio  $x_k + Z_k d_Z$ . De notar também que em ambas as definições são necessárias estimativas dos multiplicadores das restrições,  $\lambda_{k+1}$ . Biegler (1995) define  $y_k$  como

$$y_k = Z_k^T [\nabla f(x_{k+1}) - \nabla f(x_k)]$$

não sendo explicitamente necessária a obtenção de estimativas dos multiplicadores.

Para garantir que é obtida uma sequência de matrizes positivas definidas, independentemente das definições de  $d_k$  e  $y_k$  utilizadas, as aproximações à Hessiana reduzida devem ser obtidas pelo método BFGS amortecido.

### Método Gauss-Newton

Para problemas em que a função objectivo é quadrática, como frequentemente ocorre no caso de problemas de reconciliação de dados, estimativa de parâmetros e controlo predictivo, é muitas vezes razoável assumir que os valores da função objectivo e do seu gradiente são próximos de zero na solução. Nestas condições, pode demonstrar-se

que os multiplicadores também são próximos de zero e a matriz  $W_k$  pode ser simplesmente aproximada pela Hessiana da função objectivo (Biegler e Grossmann, 2004). Quando as iterações se aproximam da solução, este método é equivalente ao método de Newton, tendo uma velocidade de convergência quadrática (Dennis e Schnabel, 1983).

Considerando uma função objectivo quadrática, definida como

$$f(x) = \frac{1}{2}g(x)^T g(x)$$

e que, na solução, o valor de  $f(x^*)$  é próximo de zero, o problema (2.15) pode ser aproximado por:

$$\begin{aligned} d_k = \arg \min_d \quad & \frac{1}{2}d^T \nabla g(x_k)^T \nabla g(x_k) d + g(x_k)^T \nabla g(x_k) d \\ \text{s.a} \quad & A_g(x_k) d + c_g(x_k) = 0 \\ & A_{\mathcal{G}}(x_k) d + c_{\mathcal{G}}(x_k) \geq 0 \end{aligned}$$

Este problema corresponde à aplicação do método de Gauss-Newton ao problema (2.14). A principal vantagem deste método é o facto de não ser necessária qualquer informação de segunda ordem, quer das restrições quer da função objectivo, uma vez que a Hessiana da função objectivo pode ser obtida a partir do Jacobiano de  $g(x)$ :

$$\nabla^2 f(x_k) = \nabla g(x_k)^T \nabla g(x_k)$$

Um exemplo da aplicação desta abordagem em optimização dinâmica de processos químicos, nomeadamente em processos de polimerização, é a descrita por Silva (2005), partindo da formulação proposta por Oliveira (1994).

### 2.5.3 Métodos de globalização

Os métodos apresentados garantem a convergência para a solução do problema, desde que a estimativa inicial esteja na vizinhança da solução óptima. Quando tal não acontece, é necessário assegurar que em cada iteração há progresso em direcção à solução, relativamente à função objectivo e à satisfação das restrições. Para tal podem ser utilizados métodos de globalização alternativos, como pesquisa linear e métodos de filtro.

### Funções de mérito e pesquisa linear

Nos métodos que utilizam pesquisa linear é primeiro obtida a direcção de pesquisa (nos métodos SQP, por solução do problema (2.15), sendo depois determinado um tamanho de passo adequado nesta direcção. Para tal é necessário determinar um escalar  $\alpha_k$

$$x_{k+1} = x_k + \alpha_k d_k$$

que assegure que em cada iteração ocorre progresso razoável em direcção à solução óptima. Esse progresso é medido utilizando uma função de mérito unidimensional (pesquisa linear) que é utilizada como função objectivo num problema de optimização. No caso de pesquisa linear exacta, utilizando uma função de mérito genérica  $\phi(\alpha)$ , o tamanho do passo é a solução de

$$\alpha_k = \arg \min_{\alpha \in \mathbb{R}} \phi(\alpha) = \phi(x_k + \alpha d_k) \quad (2.29)$$

ou seja, é escolhido o valor que minimiza a função de mérito na direcção  $p_k$ , partindo de  $x_k$ .

As funções de mérito mais usuais em implementações de métodos SQP são a  $l_1$ , não diferenciável, e a Lagrangiana aumentada. A função de mérito  $l_1$  tem a forma (Fletcher, 1987)

$$\phi_1(x, \mu) = f(x) + \frac{1}{\mu} \left( \sum_{i \in \mathcal{E}} |c_i(x)| + \sum_{i \in \mathcal{I}} c_i(x)^- \right) \quad (2.30)$$

onde  $c_i(x)^- = \min(c_i(x), 0)$  e  $\mu > 0$  é o parâmetro escalar de penalidade. Para problemas apenas com restrições de igualdade (conjunto  $\mathcal{I}$  vazio), a função Lagrangiana aumentada é definida como (Nocedal e Wright, 1999)

$$\phi_F(x, \mu) = f(x) - \lambda(x)^T c(x) + \frac{1}{2\mu} \sum_{i \in \mathcal{E}} c_i(x)^2$$

onde  $\mu > 0$  é o parâmetro de penalidade e

$$\lambda(x) = [A(x)A(x)^T]^{-1} A(x)\nabla f(x)$$

são estimativas de mínimos quadrados dos multiplicadores das restrições, onde  $A^T$  é o Jacobiano das restrições. Para problemas com restrições de desigualdade esta função pode ser definida utilizando variáveis de folga. Para ambas as funções existe um escalar  $\mu^*$  tal que, para qualquer valor do parâmetro de penalidade  $0 < \mu < \mu^*$ , qualquer mínimo local do problema não linear é também um mínimo local da função de penalidade. São, por isso, funções de penalidade exactas.

A resolução do problema (2.29) pode exigir um número significativo de avaliações da função e do seu gradiente, pelo que a determinação de  $\alpha_k$  é geralmente feita por



tentativa até que seja satisfeito um determinado critério, o que resulta na chamada pesquisa linear inexacta. Um dos critérios que pode servir como base aos métodos de pesquisas linear inexacta é o constituído pelas condições de Wolfe (Nocedal e Wright, 1999)

$$\phi(\alpha_k) \leq \phi(0) + c_1 \alpha_k \phi'(0) \quad (2.31a)$$

$$\phi'(\alpha_k) \geq c_2 \phi'(0) \quad (2.31b)$$

com  $0 < c_1 < c_2 < 1$  e onde  $\phi'(0)$  é a derivada, em  $x_k$ , da função de mérito na direcção  $d_k$ . A primeira destas condições, conhecida como condição de Armijo, garante que o valor de  $\alpha$  escolhido corresponde a uma diminuição suficiente de  $\phi$ . A sua utilização isolada pode conduzir a valores de  $\alpha$  muito baixos, pelo que é necessário impor condições à curvatura, o que é feito pela segunda condição.

A verificação de ambas as condições de Wolfe implica a avaliação, para cada valor de  $\alpha$  testado, da derivada da função de mérito, o que pode ser dispendioso do ponto de vista computacional. Tal pode ser evitado se forem utilizadas estratégias de pesquisa do tipo *backtracking*. Nestas estratégias parte-se de um estimativa inicial de  $\alpha$  que vai sendo reduzida até que o critério (2.31a) seja satisfeito. Os novos valores de  $\alpha$  podem ser obtidos como o valor que minimiza um interpolante quadrático ou cúbico de  $\phi$  nos últimos valores testados.

A pesquisa linear influencia fortemente a convergência para a solução, podendo limitá-la quando são escolhidos passos inadequados. Devem, por isso, ser utilizados algoritmos robustos que garantam, em cada iteração, uma diminuição suficiente de  $\phi$  num número finito, e preferencialmente pequeno, de iterações da pesquisa linear. Um destes algoritmos é o proposto por Moré e Thuente (1994). Nos métodos SQP, a estimativa inicial do passo deve ser sempre igual à unidade, de forma a que, sempre que possível, seja esse o passo utilizado. Só desta forma é possível garantir que o método, na vizinhança da solução converge quadraticamente.

### Métodos de filtro

Uma limitação apontada à utilização de funções de mérito nos métodos SQP é, para além da necessidade da escolha do parâmetro  $\mu$ , a limitação excessiva do passo obtido pelo método de Newton, quando  $\mu$  é demasiado grande. Uma alternativa à utilização de funções de mérito é a utilização de filtros, introduzidos por Fletcher e Leyffer (2002) para utilização em métodos de regiões de confiança, sendo também possível a sua utilização nos métodos de pesquisa linear (Biegler, 2010). Na exposição seguinte iremos considerar o problema (2.14) apenas com restrições de desigualdade.

A utilização do filtro é baseada na ideia de dominação. Neste contexto, o ponto  $x_k$  domina o ponto  $x_l$  se e só se  $f(x_k) \leq f(x_l)$  e  $h(x_k) \leq h(x_l)$ , com  $h(x) = \|c_i(x)^-\|$ ,  $i \in$

$\mathcal{F}$ . O filtro é definido como uma lista de pares  $(h_l, f_l)$  tal que nenhum par domine outro. Na iteração inicial, o filtro é constituído por um único par,  $\mathcal{F}_0 = \{(U, -\infty)\}$ , onde  $U$  é o limite superior de violação das restrições.

Em cada iteração, o passo é obtido resolvendo o problema quadrático:

$$d_k = \arg \min_d \frac{1}{2} d^T W_k d + \nabla f_k^T d \quad (2.32a)$$

$$\text{s.a } c_i(x_k) + \nabla c_i(x_k)^T d \geq 0 \quad (2.32b)$$

$$\|d\|_\infty \leq \Delta_k \quad (2.32c)$$

O ponto  $x_{k+1} = x_k + d_k$  só é aceite se não for dominado pelo filtro. Se este ponto for aceite, o filtro pode ser actualizado (adicionando o par  $(h(x_k), f(x_k))$  e removendo quaisquer entradas por ele dominadas) e o raio da região de confiança aumentado. Se  $x_{k+1}$  for dominado pelo filtro, o ponto é rejeitado, o raio da região de confiança é diminuído e o problema (2.32) é resolvido novamente a partir de  $x_k$ .

A utilização do filtro como único critério de aceitação de um novo ponto não é suficiente para garantir a convergência do método, sendo necessária a introdução de algumas condições adicionais (Fletcher et al., 2007). Para garantir que as iterações não convergem para um ponto não admissível ( $h(x^*) > 0$ ), a condição de aceitação perante o filtro passa a ser

$$h(x_{k+1}) \leq \beta h_l \vee f(x_{k+1}) \leq f_l - \gamma h(x_{k+1}), \quad \forall (h_l, f_l) \in \mathcal{F}_k \cup (h(x_k), f(x_k)) \quad (2.33)$$

onde  $\beta$  e  $\gamma \in ]0, 1[$  são constantes. A aplicação desta condição, por si só, também não é suficiente para garantir a convergência para pontos estacionários. É por isso necessária a introdução de uma condição de redução suficiente, quando a restrição das violações se torna pequena. Para tal é comparada a redução esperada, dada por

$$\Delta q_k = -\frac{1}{2} d_k^T W_k d_k - \nabla f_k^T d_k$$

com a redução verificada na função objectivo. Para isso, quando  $\Delta q_k > 0$ , é necessário verificar a condição

$$f(x_k) - f(x_{k+1}) \geq \sigma \Delta q_k \quad (2.34)$$

onde  $\sigma \in ]0, 1[$  é uma constante. Assim, um ponto  $x_{k+1}$ , obtido por solução de (2.32) é aceite se satisfizer ambas as condições (2.33) e (2.34).

As propriedades de convergência global dos métodos de filtro podem ser demonstradas para alguns problemas. Nomeadamente, Fletcher et al. (2002) demonstram as propriedades de convergência destes métodos no âmbito de um método SQP em que o passo é decomposto nas suas componentes normal e tangencial. Em Fletcher et al. (2007) é feito um resumo dos principais desenvolvimentos dos métodos de fil-

tro, desde a sua introdução. São referidos os resultados de convergência obtidos e a aplicação destes métodos a diversas estratégias de optimização (como os métodos de ponto interior, apresentados na Secção 2.5.5), e também à solução de equações não-lineares.

### 2.5.4 Regiões de confiança

Nos métodos de pesquisa linear, a direcção de pesquisa é independente do tamanho do passo, sendo este encontrado inicialmente e apenas depois escolhido o passo adequado nesta direcção. Ao contrário, nos métodos de região de confiança, o tamanho do passo é determinado primeiro, sendo só depois encontrada uma direcção apropriada. O tamanho do passo adequado é incluído na solução do problema quadrático por adição da restrição:

$$\|d_k\|_2 \leq \Delta_k \quad (2.35)$$

O escalar  $\Delta_k$  pode ser interpretado como uma medida da dimensão da região, em torno de  $x_k$ , na qual se considera o problema quadrático uma aproximação satisfatória do problema original, sendo a designação regiões de confiança devida a esta interpretação. Em alguns casos, como problemas particularmente não-lineares, esta classe de métodos pode apresentar melhores propriedades de convergência do que os baseados em pesquisa linear (Biegler, 2010). No âmbito da optimização dinâmica poderão ser mais robustos na presença de incerteza devida, por exemplo, à existência de ruído numérico ou do processo (Oliveira e Biegler, 1994a).

O raio da região de confiança  $\Delta$  deve ser adaptado de forma a permitir a convergência eficiente para a solução. De forma genérica, este raio é especificado, no início de cada iteração, com base no desempenho do método nas iterações anteriores, comparando a redução efectivamente conseguida na função objectivo não linear com a redução prevista pelo modelo quadrático e verificando se a solução do problema quadrático é limitada pelo raio ( $\|d\|_2 = \Delta$ ) utilizado anteriormente. Se na iteração  $k$  a redução conseguida não for satisfatória, este parâmetro é reduzido na próxima iteração ( $\Delta_{k+1} < \Delta_k$ ).

A inclusão de (2.35) directamente no problema quadrático pode levar ao surgimento de dificuldades na sua resolução. Uma dificuldade frequente é o facto desta restrição se tornar incompatível com as restrições linearizadas, não existindo qualquer solução que satisfaça ambas simultaneamente. Este problema pode ser resolvido através da abordagem  $Sl_1QP$ , em que a função objectivo do problema quadrático é uma função

de mérito  $l_1$  do problema quadrático (2.15) (Nocedal e Wright, 1999):

$$d_k = \arg \min_d \frac{1}{2} d^T W_k d + \nabla f_k^T d + \frac{1}{\mu_k} \sum_{i \in \mathcal{E}} |c_i(x_k) + \nabla c_i(x_k)^T d| \quad (2.36a)$$

$$+ \sum_{i \in \mathcal{I}} [c_i(x_k) + \nabla c_i(x_k)^T d]^- \quad (2.36b)$$

$$\text{s.a } \|d\|_\infty \leq \Delta_k \quad (2.36c)$$

Nesta abordagem, a qualidade do passo  $d_k$  é avaliada utilizando a função de mérito  $\phi_1$  (2.30), aceitando a solução de (2.36) se a redução com ele conseguida for satisfatória. Se esta não o for, é necessário resolver novamente o problema anterior, diminuindo o valor de  $\Delta_k$ . Esta abordagem tem algumas características importantes, nomeadamente o facto de, para valores  $\mu$  suficientemente grandes, os mínimos locais de  $\phi_1$  geralmente coincidirem com soluções locais de (2.14), pelo que o método tem boas propriedades de convergência global (Nocedal e Wright, 1999).

Uma dificuldade que pode surgir com este método é o chamado efeito de Maratos, em que passos que constituem progresso em direcção à solução de (2.14) são rejeitados por causarem um aumento da função de mérito  $\phi_1$ . Este efeito pode, em alguns problemas, ocorrer repetidamente, tornando a convergência para a solução extremamente lenta. Existem, no entanto, algumas técnicas que permitem ultrapassar este problema, como, por exemplo, uma correcção de segunda ordem, em que informação localmente aproximada de segunda ordem é inserida no problema (2.36).

### 2.5.5 Métodos de ponto interior

Apesar de o seu desempenho ser geralmente bom, podem surgir algumas dificuldades na aplicação dos métodos SQP. Entre essas dificuldades estão a inconsistência das restrições linearizadas e a identificação do conjunto de restrições activas na solução do problema. Os métodos de ponto interior, originalmente desenvolvidos para problemas de programação linear, surgiram como a alternativa aos métodos SQP para problemas de grande dimensão, com elevado número de restrições de desigualdade e com estruturas especiais (Nocedal e Wright, 1999).

Os métodos de ponto interior no âmbito da programação não linear foram primeiro propostos para a resolução de problemas quadráticos, nomeadamente o problema que é necessário resolver em cada iteração do métodos SQP (Vanderbei (1994); Albuquerque et al. (1999); Dennis et al. (1994); Rao et al. (1998), entre outros). Mais recentemente foram propostos métodos que lidam directamente com o NLP, constituindo uma alternativa aos métodos SQP. Entre estes métodos destacam-se os implementados nas rotinas KNITRO (Byrd et al., 2006), LOQO (Vanderbei e Shanno, 1999) e IPOPT (Wächter e Biegler, 2006). Nesta secção será discutida a aplicação de métodos de

ponto interior a NLP's com a forma genérica (2.14).

Os métodos de ponto interior partem dos métodos de barreira em que as restrições de desigualdade do NLP (2.14) são transformadas em igualdades através da introdução de variáveis de folga. Esta transformação resulta no problema (Byrd et al., 1999)

$$\min_x \quad \varphi_\mu(x) = f(x) - \mu \sum_{i \in \mathcal{I}} \ln(s_i) \quad (2.37a)$$

$$\text{s.a.} \quad c_i(x) = 0, \quad i \in \mathcal{E} \quad (2.37b)$$

$$c_i(x) - s_i = 0, \quad i \in \mathcal{I} \quad (2.37c)$$

$$s_i \geq 0, \quad i \in \mathcal{I} \quad (2.37d)$$

onde  $s_i \in \mathbb{R}$  é a variável de folga da desigualdade  $i$  e  $\mu > 0$  é o parâmetro de barreira. Facilmente se verifica que à medida que  $c_i(x)$ ,  $i \in \mathcal{I}$ , se aproxima de zero, também a variável de folga se aproxima de zero, aumentando rapidamente o termo de barreira. Desta forma, garante-se que a solução de (2.37) satisfaz estritamente as restrições do NLP. Quando  $\mu \rightarrow 0$ , sob certas condições, a solução de (2.37) converge para uma solução local do problema original (Wächter, 2002). Daqui surge a possibilidade de resolver o NLP resolvendo, aproximadamente, uma sequência de problemas barreira com valores de  $\mu$  decrescentes.

As condições de optimalidade de KKT para o problema (2.37) podem ser escritas como um sistema de equações não lineares (Byrd et al., 2006)

$$\nabla f(x) - A_{\mathcal{E}}^T(x)\lambda - A_{\mathcal{I}}^T(x)\nu = 0 \quad (2.38a)$$

$$-\mu e + S\nu = 0 \quad (2.38b)$$

$$c_i(x) = 0, \quad i \in \mathcal{E} \quad (2.38c)$$

$$c_i(x) - s_i = 0, \quad i \in \mathcal{I} \quad (2.38d)$$

onde  $A_{\mathcal{E}}$  e  $A_{\mathcal{I}}$  são, respectivamente, os Jacobianos das restrições de igualdade e de desigualdade,  $e = [1, \dots, 1]^T$ ,  $S = \text{diag}(s_i), i \in \mathcal{I}$ , e  $\lambda$  e  $\nu$  são vectores de multiplicadores de Lagrange. Também é necessário que  $s, \nu \geq 0$ . Uma vez que  $\mu > 0$ , a condição (2.38b) implica que estas desigualdades sejam satisfeitas de forma estrita, só se podendo aproximar assintoticamente de zero quando  $\mu \rightarrow 0$ . Daí a designação de métodos de ponto interior.

Os diferentes métodos desta classe diferem entre si na forma como são obtidas soluções aproximadas de (2.37) para diferentes valores de  $\mu$  e como o valor de  $\mu$  é actualizado. Uma forma de obter a solução do problema de optimização é a aplicação do método de Newton às condições de optimalidade para obter uma direcção de

pesquisa, o que resulta no sistema de equações lineares (Byrd et al., 2006)

$$\begin{bmatrix} \nabla_{xx}^2 \mathcal{L} & 0 & -A_{\mathcal{E}}^T(x) & -A_{\mathcal{G}}^T(x) \\ 0 & Z & 0 & S \\ A_{\mathcal{E}}(x) & 0 & 0 & 0 \\ A_{\mathcal{G}}(x) & -I & 0 & 0 \end{bmatrix} \begin{bmatrix} d_x \\ d_s \\ d_\lambda \\ d_\nu \end{bmatrix} = \begin{bmatrix} \nabla f(x) - A_{\mathcal{E}}^T(x)\lambda - A_{\mathcal{G}}^T(x)\nu \\ Sv - \mu e \\ c_{\mathcal{E}}(x) \\ c_{\mathcal{G}}(x) - s \end{bmatrix}$$

onde  $d$  são as direcções de pesquisa e  $\mathcal{L}$  é a função Lagrangiana:

$$\mathcal{L}(x,s,\lambda,\nu) = f(x) - \lambda^T c_{\mathcal{E}}(x) - \nu^T (c_{\mathcal{G}}(x) - s)$$

Depois de obtida a direcção, o tamanho do passo é obtido por pesquisa linear, com *backtracking*, para garantir a não negatividade de  $s$  e  $\nu$  e uma redução suficiente de uma função de mérito. Algoritmos deste tipo são implementados nas rotinas LOQO e KNITRO.

O problema (2.37) é um problema não linear que pode ser resolvido utilizando métodos SQP com regiões de confiança. No entanto, a aplicação simples destes métodos a este problema levanta alguns problemas, nomeadamente a geração de passos que violam a não negatividade das variáveis de folga e que são frequentemente limitados pelo tamanho da região de confiança. Para ultrapassar estes problemas a rotina KNITRO implementa um método SQP especificamente desenhado para o problema de barreira. Neste método, o problema quadrático resolvido em cada iteração é (Byrd et al., 2006)

$$\min_{d_x, d_s} \nabla f(x_k)^T d_x + \frac{1}{2} d_x^T \nabla_{xx}^2 \mathcal{L}(x_k, s_k, \lambda_k, \nu_k) d_x - \mu e^T S_k^{-1} d_s + \frac{1}{2} d_s^T \Sigma_k d_s \quad (2.39a)$$

$$\text{s.a. } A_{\mathcal{E}}(x_k) d_x + c_{\mathcal{E}}(x_k) = r_{\mathcal{E}} \quad (2.39b)$$

$$A_{\mathcal{G}}(x_k) d_x - d_s + c_{\mathcal{G}}(x_k) - s_k = r_{\mathcal{G}} \quad (2.39c)$$

$$\|d_x, S_k^{-1} d_s\|_2 \leq \Delta_2 \quad (2.39d)$$

$$d_s \geq -\tau s_k \quad (2.39e)$$

onde  $\Sigma_k = S_k^{-1} Z_k$ , com  $Z_k = \text{diag}(\nu_k)$ , e  $\tau \in [0,1]$  é uma constante. As principais diferenças deste problema relativamente ao gerado pela aplicação estrita da metodologia SQP são a existência da restrição (2.39e), que garante que o passo gerado não viola a não negatividade das restrições de folga, e a presença do factor de escala  $S_k^{-1}$  na restrição do tamanho da região de confiança (2.39d), que têm em conta a proximidade ao seu limite das variáveis de folga. É também de notar que as restrições linearizadas não têm que ser estritamente satisfeitas em todas as iterações ( $r_{\mathcal{E}}, r_{\mathcal{G}} \neq 0$ ), o que permite lidar com situações em que estas são inconsistentes.

A obtenção da solução exacta do problema (2.39) não é trivial, especialmente devido à presença das restrições de desigualdade. Por essa razão, no algoritmo implemen-

tado na rotina `KNITRO` apenas são calculadas soluções inexactas, dividindo o passo nas suas componentes tangencial e normal. Tal como nos métodos de regiões de confiança descritos anteriormente, o tamanho da região de confiança vai sendo actualizado conforme a concordância entre a redução prevista pelo modelo quadrático e a efectivamente verificada.

A rotina `IPOPT` oferece diferentes opções para a resolução de NLP's através de métodos de ponto interior. Entre estas opções está a possibilidade de obter a direcção de pesquisa, utilizando métodos SQP, em que a solução dos QP's é obtida por decomposição do passo, e a implementação de um método de filtro aplicado à pesquisa linear, sendo o filtro utilizado para decidir se um determinado comprimento do passo,  $\alpha$ , é aceite ou se é necessária a sua redução.





## Capítulo 3

# Análise do processo Tennessee Eastman

Sendo um dos objectivos deste trabalho a aplicação de estratégias de optimização a processos químicos, surgiu a necessidade de seleccionar um destes processos para ser utilizado como referência, para comparação do desempenho da aplicação das diferentes estratégias consideradas. O processo escolhido foi o Tennessee Eastman (TE), apresentado por [Downs e Vogel \(1993\)](#). Esta escolha foi feita devido às características do processo, nomeadamente a sua instabilidade em ciclo aberto e a acentuada não-linearidade, características que fazem com que a aplicação de algoritmos sequenciais de optimização dinâmica não seja uma tarefa trivial. Foi ainda tido em consideração o facto de este ser um processo realista, composto por várias unidades com forte interacção entre si, e a existência de resultados publicados por outros autores, que podem servir como termo de comparação para os resultados aqui alcançados.

Neste Capítulo é feita uma descrição do processo e uma revisão dos resultados obtidos por outros autores. São depois apresentados os resultados obtidos relativamente à caracterização e ao comportamento do processo.

### 3.1 Descrição do processo

O processo Tennessee Eastman (TE) é constituído por cinco unidades principais: reactor, condensador, separador líquido/vapor, *stripper* e compressor. São produzidos dois produtos principais, designados *G* e *H*, a partir de quatro reagentes, *A*, *C*, *D* e *E*. Na Figura 3.1 é apresentado um diagrama do processo.

No reactor, alimentado pela mistura gasosa de três alimentações frescas (*A*, *D* e *E*) e

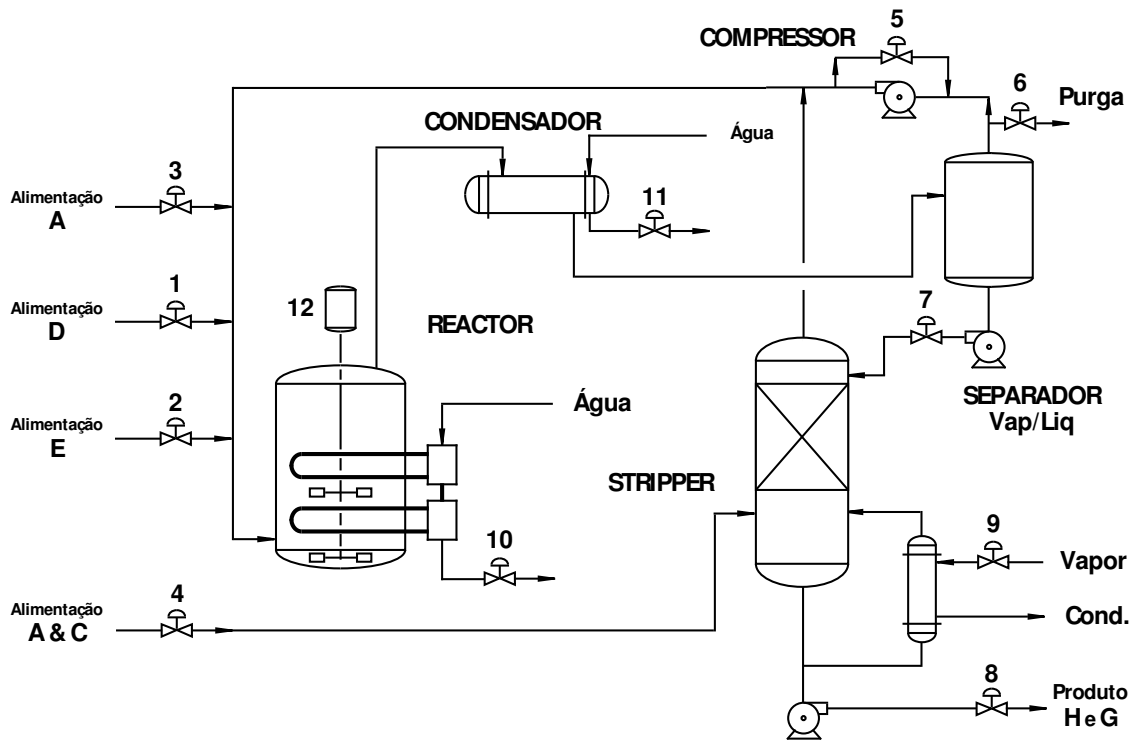


Figura 3.1 O processo Tennessee Eastman (Downs e Vogel, 1993).

uma corrente de reciclo, ocorrem quatro reacções



onde, para além dos produtos  $G$  e  $H$ , é também formado um subproduto  $F$ . Estas reacções ocorrem na presença de um catalisador, que se encontra dissolvido na fase líquida, e é totalmente involátil.

Todas as reacções são irreversíveis, exotérmicas e de aproximadamente primeira ordem relativamente às pressões parciais de cada um dos reagentes. A dependência das velocidades de reacção em relação à temperatura é descrita por uma expressão de Arrhenius, sendo a reacção que produz  $G$  aquela que tem maior energia de activação e, por isso, a mais sensível à temperatura.

Para além dos reagentes, produtos e catalisador, está ainda presente no processo um inerte gasoso  $B$ , totalmente incondensável, introduzido pelas correntes de alimentação dos reagentes. O reactor está equipado com uma serpentina de arrefecimento, onde circula água, e um agitador com velocidade variável. Depois de formados, os produtos vaporizam e abandonam o reactor numa corrente gasosa, juntamente com

**Tabela 3.1** Modos de operação do processo Tennessee Eastman (Downs e Vogel, 1993).

Modo	Razão mássica G/H	Caudal do produto	
1	50/50	7038 kg/h <i>G</i>	e 7038 kg/h <i>H</i>
2	10/90	1408 kg/h <i>G</i>	e 12669 kg/h <i>H</i>
3	90/10	10000 kg/h <i>G</i>	e 1111 kg/h <i>H</i>
4	50/50	máximo	
5	10/90	máximo	
6	90/10	máximo	

os reagentes não convertidos e o inerte.

A corrente de saída do reactor segue para o condensador onde, por arrefecimento, é parcialmente liquefeita. O condensado é maioritariamente constituído pelos produtos e por uma pequena fracção dos reagentes, sobretudo *E*. Os restantes reagentes têm calores de vaporização significativamente mais baixos que os produtos, sendo que, nas condições de operação, *A* e *C* podem ser considerados incondensáveis. O condensador tem apenas uma corrente de saída, com duas fases, que segue para o separador líquido/vapor. Depois de separadas, a fase gasosa é reciclada para o reactor através de um compressor centrífugo. Esta corrente é purgada para evitar a acumulação de inerte e subproduto no processo. A fase líquida é enviada para o *stripper*.

No *stripper* os reagentes ainda presentes na fase líquida são removidos utilizando uma alimentação fresca de *A* e *C* no estado gasoso, que circula em contra corrente com a fase líquida. A corrente gasosa, que abandona o *stripper* pelo topo, rica em reagentes, é misturada com a corrente de reciclo e enviada para o reactor. A corrente líquida, que abandona o *stripper* pela base, é maioritariamente constituída por *G* e *H* e constitui o produto final do processo. A qualidade deste produto é determinada pela razão entre os caudais mássicos de *G* e *H* nesta corrente.

No processo TE são considerados seis modos de operação, que se distinguem quer pelo caudal quer pela composição do produto. Estes modos de operação são apresentados na Tabela 3.1. O modo de operação 1 é o modo de operação base. As composições e temperaturas das correntes do processo no modo de operação base são apresentadas na Tabela 3.2, onde também são apresentadas as condições de operação das diferentes unidades e os consumos de utilidades (águas de arrefecimento do reactor e do separador, e vapor do *stripper*).

No processo TE são medidas 41 variáveis e existem 12 variáveis manipuláveis. As medições de caudais, temperaturas e pressões são contínuas, enquanto que as composições das correntes são medidas de forma descontínua. As correntes cuja composição é medida são a alimentação do reactor, a purga e o produto. Para além de serem medidas descontinuamente, estas composições são conhecidas com atraso, uma vez que

**Tabela 3.2** Balanço mássico e energético do processo TE, correspondente à operação no caso base (Downs e Vogel, 1993).

<b>Corrente</b>		Alim. A	Alim. D	Alim. E	Alim. C	Strip. topo	Alim. Rea.
Caudal molar (kgmol/h)		11.2	114.5	98.0	417.5	465.7	1890.8
Caudal mássico (kg/h)		22.4	3664.0	4509.3	6419.4	8979.6	48015.4
Temperatura (°C)		45.0	45.0	45.0	45.0	65.7	86.1
Fracções molares	A	0.99990	0.00000	0.00000	0.48500	0.43263	0.32188
	B	0.00010	0.00010	0.00000	0.00500	0.00444	0.08893
	C	0.00000	0.00000	0.00000	0.51000	0.45264	0.26383
	D	0.00000	0.99990	0.00000	0.00000	0.00116	0.06882
	E	0.00000	0.00000	0.99990	0.00000	0.07256	0.18776
	F	0.00000	0.00000	0.00010	0.00000	0.00885	0.01657
	G	0.00000	0.00000	0.00000	0.00000	0.01964	0.03561
	H	0.00000	0.00000	0.00000	0.00000	0.00808	0.01659
<b>Corrente</b>		Reactor	Reciclo	Purga	Líqu. Sep.	Produto	
Caudal molar (kgmol/h)		1476.0	1201.5	15.1	259.5	211.3	
Caudal mássico (kg/h)		48015.4	30840.0	386.5	16788.9	14288.6	
Temperatura (°C)		120.4	102.9	80.1	80.1	65.7	
Fracções molares	A	0.27164	0.32958	0.32958	0.00000	0.00479	
	B	0.11393	0.13823	0.13823	0.00000	0.00009	
	C	0.19762	0.23978	0.23978	0.00000	0.01008	
	D	0.01075	0.01257	0.01257	0.00222	0.00018	
	E	0.17722	0.18559	0.18579	0.13704	0.00836	
	F	0.02519	0.02263	0.02263	0.01669	0.00099	
	G	0.12302	0.04844	0.04844	0.47269	0.53724	
	H	0.08423	0.02299	0.02299	0.37136	0.43828	
<b>Condições de operação</b>							
		Reactor	Separador	Condens.	Stripper		
Temperatura (°C)		120.4	80.1	—	65.7		
Pressão (kPa rel.)		2705.0	2633.7	—	3102.2		
Calor transferido (kW)		-6468.7	—	-2140.6	1430.0		
Volume de líquido (m <sup>3</sup> )		16.55	4.88	—	4.43		
<b>Utilidades</b>							
Água de arrefecimento do reactor (m <sup>3</sup> /h)				93.37			
Água de arrefecimento do condensador (m <sup>3</sup> /h)				49.37			
Vapor do stripper (kg/h)				230.31			

o processo de análise não pode ser considerado instantâneo. As variáveis medidas continuamente e as variáveis manipuláveis são apresentadas nas Tabelas 3.3 e 3.4, respectivamente. Os valores apresentados correspondem ao modo base de operação.

Os caudais das correntes de reciclo do compressor e da corrente de purga não dependem exclusivamente da posição das válvulas  $u_5$  e  $u_6$ , mas também das diferenças

**Tabela 3.3** Variáveis do processo TE medidas continuamente (Downs e Vogel, 1993).

Nome	Número	Caso base	Unidades
Caudal da alimentação de A	$y_1$	0.25052	kscmh <sup>a</sup>
Caudal da alimentação de D	$y_2$	3664.0	kg/h
Caudal da alimentação de E	$y_3$	4509.3	kg/h
Caudal da alimentação de C	$y_4$	9.3477	kscmh <sup>a</sup>
Caudal do reciclo	$y_5$	26.902	kscmh <sup>a</sup>
Caudal da alimentação reactor	$y_6$	42.339	kscmh <sup>a</sup>
Pressão do reactor	$y_7$	2705.0	kPa rel.
Nível do reactor	$y_8$	75.000	%
Temperatura do reactor	$y_9$	120.40	°C
Caudal da purga	$y_{10}$	0.33712	kscmh <sup>a</sup>
Temperatura do separador	$y_{11}$	80.109	°C
Nível do separador	$y_{12}$	50.000	%
Pressão do separador	$y_{13}$	2633.7	kPa rel.
Caudal da base do separador	$y_{14}$	25.160	m <sup>3</sup> /h
Nível do <i>stripper</i>	$y_{15}$	50.000	%
Pressão do <i>stripper</i>	$y_{16}$	3102.2	kPa rel.
Caudal do produto	$y_{17}$	22.949	m <sup>3</sup> /h
Temperatura do <i>stripper</i>	$y_{18}$	65.731	°C
Caudal de vapor do <i>stripper</i>	$y_{19}$	230.31	kg/h
Potência do compressor	$y_{20}$	341.43	kW
Temp. de saída da água de arref. do reactor	$y_{21}$	94.599	°C
Temp. de saída da água de arref. do sep.	$y_{22}$	77.297	°C

<sup>a</sup> Thousand standard cubic meters per hour.

de pressão entre as unidades. Assim, podem ocorrer variações destes caudais sem que ocorra qualquer movimento das válvulas. A velocidade de agitação do reactor ( $u_{12}$ ) apenas influencia o coeficiente de transferência de calor entre a fase líquida do reactor e a serpentina de arrefecimento, não tendo qualquer influência no grau de mistura (que se supõe sempre completa), nem sobre as velocidades de reacção. Todas as variáveis manipuláveis podem tomar qualquer valor entre 0% e 100% do intervalo entre os limites inferior e superior.

É também importante notar que o caudal de vapor no *stripper*, variável  $y_{19}$ , não é medido directamente. O seu valor é obtido supondo que existe equilíbrio térmico entre o vapor e o *stripper* e que a temperatura do vapor ser mantém constante. A partir do valor da posição da válvula que controla o caudal de vapor ( $u_9$ ), é determinado o coeficiente de transferência de calor. Com este valor e com a temperatura do *stripper* ( $y_{18}$ ), utilizando a equação de transferência de calor por condução em equilíbrio, é calculada a quantidade de calor transferida. Conhecida esta quantidade, determina-se o caudal de vapor. Uma consequência deste facto é a possível mudança no caudal de

**Tabela 3.4** Variáveis manipuláveis do processo TE (Downs e Vogel, 1993).

Nome	Número	Caso base (%)	Limites		Unidades
			inf.	superior	
Caudal de D	$u_1$	63.053	0	5881.	kg/h
Caudal de E	$u_2$	53.980	0	8354.	kg/h
Caudal de A	$u_3$	24.644	0	1.1017	kscmh
Caudal de C	$u_4$	61.302	0	15.25	kscmh
Válvula de reciclo do compressor	$u_5$	22.210	0	100.	%
Válvula da purga	$u_6$	40.064	0	100.	%
Caudal de líquido do separador	$u_7$	38.100	0	65.71	m <sup>3</sup> /h
Caudal de produto	$u_8$	46.534	0	49.10	m <sup>3</sup> /h
Válvula de vapor do <i>stripper</i>	$u_9$	47.446	0	100.	%
Caudal da água de arrefecimento do rector	$u_{10}$	41.106	0	227.1	m <sup>3</sup> /h
Caudal da água de arrefecimento do condensador	$u_{11}$	18.114	0	272.6	m <sup>3</sup> /h
Velocidade do agitador	$u_{12}$	50.000	150	250.	rpm

vapor medido, sem que haja qualquer alteração da posição da válvula que o controla.

Na Tabela 3.5 são apresentadas as variáveis do processo medidas descontinuamente. São também apresentados os intervalos de amostragem e os atrasos com que as medições são conhecidas. Os valores apresentados são relativos ao modo base de operação.

Os intervalos de amostragem e os tempos mortos são características de cada um dos analisadores, não tendo qualquer relação com a estrutura de controlo. No caso da corrente de alimentação do rector, tanto o intervalo de amostragem como o tempo morto são iguais a 0.1 h, ou seja, novos valores da composição desta corrente estão disponíveis de 0.1 em 0.1h, e correspondem à composição do caudal 0.1h antes.

### 3.1.1 Modelo do processo

O modelo do processo é fornecido por Downs e Vogel (1993) na forma de um conjunto de rotinas em FORTRAN, que permitem simular o seu comportamento dinâmico. Não é fornecido qualquer modelo analítico. Este modelo contém 50 variáveis de estado, 41 variáveis medidas e 12 variáveis manipuláveis. As equações diferenciais que fazem parte do modelo são:

- balanços mássicos parciais ao rector (oito);
- balanços mássicos parciais ao separador (oito);
- balanços mássicos parciais ao *stripper* (oito);

**Tabela 3.5** Variáveis medidas descontinuamente no processo TE (Downs e Vogel, 1993).

Análise da alimentação do reactor				
Composto	Número	Caso base	Unidades	Intervalo de amostragem = 0.1 h Tempo morto = 0.1 h
A	$y_{23}$	31.188	mol%	
B	$y_{24}$	8.8933	mol%	
C	$y_{25}$	26.383	mol%	
D	$y_{26}$	6.8820	mol%	
E	$y_{27}$	18.776	mol%	
F	$y_{28}$	1.6567	mol%	
Análise da purga				
Composto	Número	Caso base	Unidades	Intervalo de amostragem = 0.1 h Tempo morto = 0.1 h
A	$y_{29}$	32.958	mol%	
B	$y_{30}$	13.823	mol%	
C	$y_{31}$	23.978	mol%	
D	$y_{32}$	1.2565	mol%	
E	$y_{33}$	18.579	mol%	
F	$y_{34}$	2.2633	mol%	
G	$y_{35}$	4.8436	mol%	
H	$y_{36}$	2.2986	mol%	
Análise do produto				
Composto	Número	Caso base	Unidades	Intervalo de amostragem = 0.25 h Tempo morto = 0.25 h
D	$y_{37}$	0.01787	mol%	
E	$y_{38}$	0.83570	mol%	
F	$y_{39}$	0.09858	mol%	
G	$y_{40}$	53.724	mol%	
H	$y_{41}$	43.828	mol%	

- balanços mássicos parciais à zona de mistura de alimentação do reactor (oito);
- balanço energético ao reactor;
- balanço energético ao separador;
- balanço energético ao *stripper*;
- balanço energético à zona de mistura da alimentação do reactor;
- balanço energético à água de arrefecimento do reactor;

- balanço energético à água de arrefecimento do separador;
- dinâmica das variáveis manipuláveis (doze).

Na elaboração do modelo foi considerado que todos os gases se comportam como gases ideais, que todas as fases líquidas são perfeitamente agitadas e que o equilíbrio líquido vapor segue a lei de Raoult, sendo a pressão de vapor calculada pela equação de Antoine. De forma a evitar uma complexidade acrescida do modelo, supõe-se também que, quando em contacto, as fases líquida e gasosa se encontram em equilíbrio e que este equilíbrio se estabelece instantaneamente.

Uma característica importante deste modelo é sua forte não-linearidade. Para além das cinéticas das reacções e das equações de equilíbrio, contribui também para esta não-linearidade o facto de ser considerada a variação, com a temperatura, das capacidades caloríficas das espécies químicas presentes. Assim, os balanços energéticos às diferentes unidades são feitos em função da entalpia total. O valor da temperatura é obtido a partir da entalpia pela solução numérica de uma equação não-linear, utilizando o método de Newton. De notar também que se considera a entalpia total de cada uma das unidades igual à entalpia da fase líquida, aproximação que se justifica pela proporções entre as quantidades desta fase e da fase gasosa. Com esta simplificação pretende-se evitar a necessidade de realizar cálculos relativos ao equilíbrio líquido/vapor em cada iteração do método de Newton no cálculo da temperatura.

### 3.1.2 Objectivos de controlo

De acordo com [Downs e Vogel \(1993\)](#), os principais objectivos de controlo para o processo TE são:

- manter as variáveis do processo próximas dos valores desejados;
- manter as condições de operação dentro dos limites de segurança dos equipamentos;
- minimizar as variações no caudal e na qualidade do produto;
- minimizar o movimento das válvulas, especialmente das alimentações frescas;
- recuperar rápida e suavemente de perturbações inesperadas;
- alterar o modo de operação de forma rápida e suave.

As perturbações inesperadas mais frequentes neste processo estão incluídas nas rotinas de simulação. Algumas destas perturbações têm um carácter aleatório e outras são totalmente desconhecidas. Todas as perturbações podem ser iniciadas e interrompidas em qualquer altura, quer isoladamente quer em conjunto.



**Tabela 3.6** Limites normais de operação e de segurança do processo Tennessee Eastman (Downs e Vogel, 1993).

Variável	Limites de operação		Limites de segurança	
	Inferior	Superior	Inferior	Superior
Pressão do reactor		2895 kPa		3000 kPa
Nível do reactor	50% (11.8 m <sup>3</sup> )	100% (21.3 m <sup>3</sup> )	2.0 m <sup>3</sup>	24.0 m <sup>3</sup>
Temp. do reactor		150 °C		175 °C
Nível no separador	30% (3.3 m <sup>3</sup> )	100% (9.0 m <sup>3</sup> )	1.0 m <sup>3</sup>	12.0 m <sup>3</sup>
Nível do <i>stripper</i>	30% (3.5 m <sup>3</sup> )	100% (6.6 m <sup>3</sup> )	1.0 m <sup>3</sup>	8.0 m <sup>3</sup>

No processo TE, por razões de segurança, existem limites rígidos para as condições de operação de algumas unidades, cuja violação implica a paragem imediata do processo. É, por isso, fundamental que o processo seja mantido dentro destes limites, apresentados na Tabela 3.6.

A amplitude de variação permitida é, para algumas variáveis, bastante pequena. Por exemplo, os limites de operação e segurança para a pressão do reactor estão, respectivamente, 7% e 11% acima do seu valor no modo base de operação.

O processo TE está inserido numa unidade industrial onde, a montante, são produzidos os reagentes utilizados e, a jusante, os produtos *G* e *H* são separados e purificados. Ainda que exista alguma capacidade de armazenamento intermédio, esta é limitada, pelo que devem ser consideradas as variações possíveis nos caudais das alimentações frescas e avaliadas as consequências da variação na composição e caudal do produto. Para assegurar o correcto funcionamento do processo a jusante, devem ser evitadas variações no caudal de produto superiores a  $\pm 5\%$  com conteúdo significativo de frequência no intervalo 8-16 h<sup>-1</sup>. A composição de *G* (%mol) no produto não deve sofrer variações superiores a  $\pm 5\%$  com conteúdo significativo de frequência no intervalo 6-10 h<sup>-1</sup>.

Relativamente ao fornecimento das alimentações frescas dos reagentes, não existem restrições significativas quanto à alimentação de *E*, uma vez que a capacidade de armazenagem disponível é significativa. Para as alimentações de *A* e *D*, as capacidades de armazenagem são mais reduzidas, sendo extremamente limitada para a alimentação de *C*. Assim, devem ser evitadas variações no caudal de alimentação de *C*, sobretudo com conteúdo significativo de frequências na gama 12-80 h<sup>-1</sup>. Para as alimentações de *A* e *D* devem ser evitadas variações com conteúdo significativo de frequências na gama 8-16 h<sup>-1</sup>. A variação do caudal de alimentação de *E* não suscita preocupações, mas deve ser evitado o movimento excessivo das válvulas.

## 3.2 Resultados publicados por outros autores

Desde a sua publicação, o processo da Tennessee Eastman foi utilizado por vários autores para testar e comparar diversas ferramentas. Nesta secção são apresentados e discutidos os resultados publicados resultantes da aplicação de ferramentas de identificação de processos, controlo e optimização a este processo.

As estruturas de controlo apresentadas tentam satisfazer os objectivos propostos por [Downs e Vogel \(1993\)](#). Estas estruturas podem, de forma genérica, ser divididas em controlo baseado em modelos e estruturas de controlo descentralizado. As estruturas de controlo descentralizado são constituídas por um conjunto de controladores SISO, seleccionados e sintonizados de forma a conseguir um desempenho adequado de todo o processo. Uma das vantagens geralmente apontadas a este tipo de estrutura é o facto de não requererem um modelo dinâmico do processo. Nesta categoria podem ser incluídos os resultados apresentados por [Price et al. \(1994\)](#); [McAvoy e Ye \(1994\)](#); [Vinson et al. \(1995\)](#); [Lyman e Georgakis \(1995\)](#); [Banerjee e Arkun \(1995\)](#); [Ricker \(1996\)](#); [Luyben \(1996\)](#); [Luyben et al. \(1997\)](#); [McAvoy \(1998, 1999\)](#); [Larsson et al. \(2001\)](#); [Antelo et al. \(2008\)](#). Alguns destes trabalhos são discutidos e comparados por [Larsson e Skogestad \(2000\)](#) que, para além desta discussão, fazem um resumo das metodologias de projecto para este tipo de estrutura de controlo.

Este processo também foi utilizado para testar diferentes metodologias de controlo que utilizam o modelo do processo para prever o seu comportamento, de forma a otimizar o seu desempenho. Este tipo de estratégia foi utilizado por [Ricker \(1993\)](#); [Ricker e Lee \(1995b,a\)](#); [Srinivas e Arkun \(1997\)](#); [Zheng \(1998\)](#); [Becerra et al. \(1999\)](#); [McAvoy \(2002\)](#); [Jockenhövel et al. \(2003\)](#); [Tian e Hoo \(2005\)](#), sendo que os resultados apresentados por [Ricker \(1993\)](#) são apenas para o reactor.

Os trabalhos até aqui referidos propõem estruturas de controlo para atingir os objectivos propostos por [Downs e Vogel \(1993\)](#), operando em torno dos diferentes modos de operação. Para cada um destes seis modos de operação, [Ricker \(1995\)](#) obtém condições de operação economicamente mais favoráveis, de acordo com a função custo fornecida. Também de forma a obter condições de operação mais favoráveis, [Duvall e Riggs \(2000\)](#) e [Golshan et al. \(2005\)](#) aplicam algoritmos de optimização em tempo real ao processo. O problema da determinação das condições óptimas de operação também é abordado por [Blanco et al. \(2004\)](#), considerando adicionalmente um critério de operacionalidade, relacionado com a estabilização do processo em ciclo aberto. Não tendo em consideração questões económicas, [Kavouras et al. \(2005\)](#) aplicam ao processo TE uma metodologia para a determinação de estados estacionários a partir de modelos do tipo caixa negra, aplicável a processos instáveis.

São ainda de referir os trabalhos de [Juricek et al. \(2001\)](#); [Kraus et al. \(2006\)](#), ambos no âmbito da identificação e estimativa de estados, e de [Zerkaoui et al. \(2010\)](#),

resultante da aplicação de redes neuronais ao controlo do processo.

[Price et al. \(1994\)](#) propõem um conjunto de regras qualitativas para a selecção e emparelhamento das variáveis, com o objectivo fundamental de assegurar um controlo adequado do caudal de produto, e também das quantidades das diferentes espécies químicas envolvidas no processo. De entre os candidatos ao controlo do caudal de produto, devem ser preferidos os caudais “internos” do processo. Só depois de satisfeitos estes objectivos, são consideradas as questões relacionadas com a qualidade do produto. Utilizando estas regras, são identificadas quatro estruturas de controlo descentralizado para todo o processo, que são testadas por simulação do processo. Uma das estruturas, em que o caudal de produto é controlado utilizando o caudal de água de arrefecimento do separador (um caudal “interno” do processo), demonstra ter um bom desempenho em torno do modo base de operação. Resultados semelhantes são apresentados por [Lyman e Georgakis \(1995\)](#), que também comparam o desempenho das quatro estruturas utilizando o custo de operação.

[McAvoy e Ye \(1994\)](#) utilizam critérios quantitativos, como a matriz de ganhos relativos (*relative gain array*, RGA), para a selecção das variáveis e para o projecto da estrutura de controlo, feita em três etapas. Na primeira etapa são fechados todos os ciclos de controlo internos, controlando todos os caudais medidos e as temperaturas do reactor e do separador, sendo os *set-points* destes ciclos variáveis manipuláveis. Esta estrutura é projectada para operação apenas em torno do modo base de operação.

[Ricker \(1996\)](#) propõe uma abordagem diferente ao problema de controlo do processo, baseada em heurísticas e informação qualitativa sobre o comportamento dinâmico do processo. Esta abordagem pretende ultrapassar as dificuldades verificadas nas soluções anteriormente propostas em lidar com as restrições impostas às condições de operação, sendo este aspecto de extrema importância quando se pretende maximizar a produção, ou minimizar os custos, situações que invariavelmente resultam na existência de variáveis nos seus limites. A principal característica diferenciadora da estrutura de controlo apresentada é o facto de os *set-points* de todos os ciclos de controlo de caudais serem estabelecidos com base no caudal e composição pretendidos para o produto, sendo que o controlo da composição é feito recorrendo a uma componente *feedforward*.

Os resultados apresentados demonstram que este esquema permite obter bons resultados para todos os modos de operação do processo, sendo mesmo semelhantes aos apresentados por [Ricker e Lee \(1995a\)](#), utilizando controlo predictivo. Em algumas situações, como quando existem variáveis nos seus limites, o autor considera o desempenho da estrutura descentralizada superior à do controlador predictivo, o que resulta da forma como os algoritmos de optimização resolvem o problema. Ainda que a abordagem utilizada seja sobretudo qualitativa, algumas das decisões são tomadas com base em resultados de simulação utilizando o modelo dinâmico do processo. O autor refere que, para além do conhecimento profundo das características do pro-

cesso, a existência deste modelo é essencial, por permitir a realização de testes que dificilmente seriam possíveis numa instalação real.

Luyben et al. (1997) propõem um procedimento heurístico para o projecto de estruturas de controlo descentralizado. A aplicação deste procedimento ao processo TE é detalhadamente descrito por Luyben (1996). De forma simples o procedimento pode ser dividido em cinco partes: estabelecer como o caudal de produto é controlado; seleccionar as variáveis que permitam um melhor controlo da qualidade e das restrições; seleccionar variáveis para controlar os níveis e as pressões; assegurar que os balanços mássicos globais são satisfeitos, particularmente para os inertes; utilizar as restantes variáveis manipuláveis para otimizar critérios económicos ou o comportamento dinâmico. A estrutura resultante é simples e as razões das escolhas feitas são claras.

Uma característica da estrutura utilizada é o facto do caudal de produto ser directamente controlado pela válvula que controla o caudal de saída da base do *stripper*, tendo os restantes controladores de nível que acomodar esta escolha. No entanto, os testes dinâmicos realizados mostram que esta estrutura não é suficiente para satisfazer os objectivos de controlo propostos por Downs e Vogel (1993). Para tal é necessária a utilização de controladores adicionais (*overrides*), seleccionados de acordo com as dificuldades observadas na simulação. Todos os controladores SISO utilizados são proporcionais. Os resultados apresentados demonstram que esta estrutura é capaz de satisfazer aqueles que, segundo o autor, são os principais objectivos de controlo.

McAvoy (1998) utiliza o processo para demonstrar uma metodologia de selecção de controladores de nível em estruturas de controlo descentralizado. O autor começa por analisar as características do modelo, utilizando uma versão linearizada em torno do modo base de operação. Este autor refere ainda a existência de dois valores singulares nulos da matriz de estado do modelo linear, bem como de dois valores próprios nulos e quatro valores próprios com parte real positiva na mesma matriz. Estes valores indicam a existência de modos instáveis e de dois modos integrantes, que o autor identifica como sendo os níveis do separador e do *stripper*. É demonstrado que o nível do reactor não é integrante. São depois comparadas quinze diferentes estratégias de controlo dos níveis integrantes, concluindo-se que apenas duas delas são viáveis, apresentando as restantes problemas de estabilidade. A sua utilização, nomeadamente aquela em que são utilizados caudais das alimentações, pode resultar na alteração do sinal dos ganhos estacionários, dificultando a tarefa de controlo do processo com controladores lineares.

McAvoy (1999) propõe uma estratégia de síntese de controlo para todo o processo baseada na resolução de problemas de optimização MILP, utilizando como critério a minimização da variação das variáveis manipuladas necessária para que as variáveis medidas retornem ao seu valor em estado estacionário após uma perturbação. É utilizado um modelo de ganho, do tipo  $y = Ku$ , obtido utilizando a metodologia de

McAvoy e Miller (1999) para sistemas que contenham variáveis integrantes e para os quais esteja disponível um modelo dinâmico. O problema de otimização resultante não é resolvido globalmente para todo o processo, sendo primeiro obtida uma solução para as variáveis medidas mais importantes do ponto de vista da segurança do processo, depois para as variáveis relacionadas com a quantidade e qualidade do produto e, finalmente, para as restantes variáveis. As soluções obtidas para cada um dos conjuntos de variáveis são incluídas no modelo e consideradas na solução dos problemas seguintes. O problema de otimização é resolvido por enumeração de todas as soluções possíveis, que são ordenadas de acordo com o critério utilizado. Uma dificuldade importante é a necessidade de resolver um problema para cada uma das diferentes perturbações consideradas, podendo ser obtidas diferentes soluções para diferentes perturbações. Por essa razão são ainda utilizadas outras ferramentas de selecção, como a matriz dos ganhos relativos e o índice de Niederlinsky. A sintonização dos controladores SISO é feita *a posteriori*.

Larsson et al. (2001) apresentam um procedimento sistemático para a obtenção de uma estrutura de controlo designada *self-optimizing*, em que se pretende identificar quais as variáveis que, quando mantidas constantes na presença de perturbações, resultam numa perda económica aceitável. Estas variáveis devem ser controladas utilizando *set-points* constantes. Este é um problema complexo, mesmo no caso do processo TE em que, após análise dos graus de liberdade, se verifica que o número de variáveis que é necessário escolher é igual a três. Este valor resulta da necessidade de controlar os níveis do separador e do *stripper*, de satisfazer a composição e caudal do produto, e do número de restrições activas nas condições de custo mínimo para o modo base de operação (cinco, segundo os resultados apresentados por Ricker (1995) que são utilizados pelos autores). São apresentadas algumas regras que, aplicadas ao processo TE permitem reduzir o elevado número de alternativas possíveis a analisar. Após aplicação destas regras, o número de alternativas a considerar é de 165 possíveis combinações de três variáveis. A análise das perdas para estas combinações permite reduzir mais uma vez o número de combinações a considerar, resultando em dois grupos de três variáveis (temperatura do reactor, C na purga e caudal de reciclo; temperatura do reactor, C na purga e trabalho do compressor).

Após esta análise é necessário definir os emparelhamentos entre as variáveis de saída e as variáveis de entrada, bem como sintonizar os controladores SISO a utilizar. Depois de completada esta tarefa, as diferentes alternativas podem ser simuladas, sendo escolhida aquela com que for conseguido melhor desempenho. Os autores referem que a solução obtida pelo procedimento proposto pode não ser óbvia, e ser substancialmente diferente da obtida utilizando outros critérios mais usuais. É também de notar que este procedimento apenas permite identificar quais as variáveis que devem ser controladas. A forma como são controladas é escolhida de forma semelhante à utilizada noutros trabalhos aqui apresentados. Toda a análise, nomeadamente a avaliação do número de graus de liberdade, é feita para as condições correspondentes ao

custo mínimo de operação no modo base de operação.

[Antelo et al. \(2008\)](#) utilizam princípios termodinâmicos e a representação do processo como uma rede onde é transferida massa e energia entre os diferentes nós. De forma a simplificar o projecto da estrutura de controlo o processo é decomposto em duas subredes, uma de reacção e outra de separação. Cada uma destas subredes é estabilizada separadamente. Ainda que a metodologia utilizada estabilize cada uma das subredes consideradas, não estabiliza globalmente o processo, nomeadamente devido à dificuldade em controlar a temperatura no reactor, devido à saturação da válvula do caudal de arrefecimento. Pela introdução de alterações na metodologia anteriormente considerada, este problema é resolvido e é obtida uma estrutura capaz de estabilizar o processo.

Os primeiros resultados da aplicação de controlo predictivo ao processo da Tennessee Eastman foram publicados por [Ricker \(1993\)](#) e referem-se a uma versão simplificada do reactor do processo, motivado pelas dificuldades observadas na obtenção de uma estratégia de controlo predictivo para todo o processo. Essas dificuldades devem-se à existência de limites rígidos de segurança e à degradação do comportamento do processo para algumas perturbações. Ao contrário do modelo original, o modelo considerado é estável em ciclo aberto, o que resulta da inclusão de um controlador SISO proporcional que controla o nível do reactor manipulando o caudal de produto. No algoritmo de controlo predictivo é utilizado um modelo discreto, linear e invariante no tempo, sendo apenas considerado um subconjunto das variáveis medidas disponíveis. Os resultados obtidos não são satisfatórios, o que, segundo o autor, se deve às variações dos ganhos relativos. Esta variação, associada à forte interacção entre as variáveis do processo, pode levar à degradação do comportamento do sistema quando é utilizado o controlador. Para resolver este problema é testado um controlador predictivo descentralizado, constituído por vários controladores em que apenas são considerados subconjuntos das variáveis medidas e manipuláveis. Também esta estratégia não consegue controlar adequadamente o reactor. O autor sugere que, dadas as suas características, a forma mais viável de resolver este problema será a utilização de controladores não-lineares, naturalmente capazes de lidar com as variações dos ganhos.

[Ricker e Lee \(1995b\)](#) apresentam um modelo dinâmico, mecanístico e não-linear do processo que, segundo os autores, descreve as características dinâmicas essenciais do processo, não sendo demasiado complexo. Uma particularidade importante deste modelo é ser considerada a existência *a priori* de dois controladores PI para as temperaturas do reactor e do separador, que permitem converter estas temperaturas em variáveis manipuláveis (através dos *set-points* dos respectivos controladores). Assim, não são necessários balanços energéticos a estas unidades e, conseqüentemente, não é necessário modelar os fenómenos de geração e transferência de energia. O modelo resultante contém 26 variáveis de estado, 10 variáveis manipuláveis, 23 variáveis medidas e 123 parâmetros. Destes, 15 são ajustados *on-line* utilizando um filtro de



Kalman generalizado (*extended Kalman filter*). A escolha do conjunto de parâmetros a ajustar *on-line* é feita de forma a conseguir que o modelo produza estimativas, sem desvios, do mesmo número de variáveis medidas, variáveis estas que, por sua vez, são escolhidas utilizando argumentos qualitativos (por exemplo, variáveis para as quais é importante a não ocorrência de *offset*). O filtro de Kalman generalizado também é utilizado para estimar as variáveis de estado do modelo, que por sua vez são utilizados para prever as variáveis medidas em instantes futuros, através da integração do modelo. São apresentados resultados de simulação que mostram uma boa concordância entre o processo e o modelo apresentado.

Utilizando o modelo descrito, [Ricker e Lee \(1995a\)](#) propõem uma estratégia de controlo predictivo não-linear para todo o processo da Tennessee Eastman. O modelo do processo utilizado no algoritmo de optimização é obtido por diferenciação numérica em torno das variáveis de estado e manipuláveis correspondentes ao instante de amostragem no qual se pretende resolver o problema de controlo. Daqui resulta que a trajectória de controlo é obtida pela solução de um problema quadrático. A utilização de controlo predictivo para este processo é motivada pelos resultados obtidos utilizando estruturas de controlo descentralizado (como a proposta por [McAvoy e Ye \(1994\)](#)), cujos resultados são utilizados como termo de comparação) e por outras estratégias de controlo multivariável (como *dynamic matrix control*, DMC, *quadratic programming solution of dynamic matrix control*, QDMC), todos eles revelando algumas limitações. No caso do controlo descentralizado, as limitações identificadas são a frequente saturação dos ciclos de controlo, que torna necessária utilização de ciclos adicionais (*overrides*) difíceis de projectar, e a dificuldade em lidar com a variação do ganho das saídas relativamente às entradas. Esta última também afecta o desempenho das estratégias de controlo multivariável referidas. De acordo com os autores, as variáveis mais problemáticas são a pressão do reactor e os níveis, sendo referido que o ganho da pressão do reactor com o caudal da alimentação de A muda de sinal na vizinhança do estado estacionário correspondente ao modo base de operação.

Antes da aplicação do algoritmo de controlo predictivo, o processo é estabilizado por um conjunto de cinco controladores SISO PI, que acrescem aos dois controladores SISO já considerados no modelo. São também considerados controladores SISO para todas as alimentações e a purga, que manipulam as posições das respectivas válvulas. O controlador predictivo utiliza 8 variáveis medidas e 8 variáveis manipuladas. Estas variáveis são escolhidas de forma semelhante à utilizada por [McAvoy e Ye \(1994\)](#) para o emparelhamento com controladores SISO. Este número reduzido de variáveis é possível devido à utilização dos referidos controladores SISO, cujos *set-points* são ajustados pelo controlador predictivo. Os resultados obtidos são, em todas as situações, superiores aos obtidos por estratégias de controlo descentralizado, sendo que as maiores diferenças ocorrem em situações em que as condições do processo se aproximam dos limites de segurança e quando são feitas mudanças de *set-point*. As conclusões apresentadas apontam no sentido de ser possível melhorar o desempenho utilizando

um modelo não-linear em todos os passos do algoritmo de controlo.

Sriniwas e Arkun (1997) aplicam o algoritmo de controlo QDMC (Qin e Badgwell, 2003), utilizando um modelo linear do processo, do tipo *input output*, em que são consideradas quatro variáveis medidas e quatro variáveis manipuláveis. Este modelo é obtido considerando a estrutura de controlo descentralizado proposta por Banerjee e Arkun (1995), sendo as quatro variáveis manipuláveis do controlador predictivo *set-points* de quatro dos controladores SISO. O modelo é obtido utilizando técnicas de identificação. Os dados utilizados são submetidos a um teste de linearidade, verificando os autores a existência de uma relação linear entre as entradas e as saídas consideradas. Esta relação linear é devida à existência dos controladores SISO que atenuam a não-linearidade do processo. Os resultados obtidos são satisfatórios, ainda que o desempenho não seja tão bom como o conseguido com controlo predictivo por Ricker e Lee (1995a).

Zheng (1998) aplica um algoritmo de controlo predictivo não-linear em que apenas são calculadas as variáveis manipuladas a implementar no primeiro intervalo de amostragem, sendo as variáveis manipuladas no restante horizonte de previsão aproximadas por outro controlador. Na aplicação ao processo TE, este controlador é um controlador predictivo linear, que utiliza uma versão linearizada do modelo de Ricker e Lee (1995b), em torno das condições de operação correspondente ao modo base. O mesmo modelo é utilizado no algoritmo de controlo que obtém as variáveis de controlo no primeiro instante de amostragem.

Becerra et al. (1999) utiliza o processo TE para testar diferentes alternativas de utilização de critérios económicos com MPC. O modelo do processo utilizado é linear (ARMAX), obtido por identificação, e considera a utilização de uma das estruturas de controlo descentralizado apresentadas em Luyben et al. (1997). O algoritmo de controlo predictivo ajusta os *set-points* de cinco dos controladores SISO que constituem esta estrutura, a partir de seis variáveis, sendo cinco delas variáveis medidas do processo e uma o custo de operação. As alternativas consideradas são a utilização de um algoritmo de optimização em estado estacionário que determina o conjunto de referências óptimas a serem enviadas ao controlador predictivo, a utilização de optimização multiobjectivo no controlador predictivo e a utilização de controlo predictivo com um *set-point* para o custo. Os autores concluem que estas duas últimas alternativas são mais favoráveis por permitirem a optimização dos custos em estado transiente.

Jockenhövel et al. (2003) apresentam resultados da aplicação de controlo predictivo não-linear ao processo, sendo que no modelo utilizado são considerados os balanços energéticos. No entanto, estes têm uma forma diferente do modelo original, considerando a temperatura como a variável diferencial a eles associada. Existem ainda outras diferenças, nomeadamente ao nível dos balanços mássicos, que resultam num modelo com 32 variáveis de estado diferenciais e 11 variáveis manipuláveis (não é



considerada a velocidade de agitação do reactor). Não é considerada a existência de qualquer estrutura de controlo de base, sendo o modelo do processo instável em ciclo aberto. Este modelo é directamente utilizado pelo controlador predictivo, sendo o problema resolvido utilizando um método simultâneo. É utilizado um intervalo de amostragem de 100s, verificando os autores que a utilização de intervalos de amostragem maiores, juntamente com a instabilidade do sistema, resultam em oscilações de algumas variáveis de entrada. Na resolução do problema de optimização resultante é comparado o desempenho de dois algoritmos de optimização distintos, IPOPT (Wächter, 2002) e SNOPT (Gill et al., 1997a). Verifica-se que o desempenho deste último não é satisfatório, devido à demorada convergência ou até, em alguns casos, à não convergência no limite de 100 iterações. São apresentados resultados para diversas mudanças de *set-point*.

Tian e Hoo (2005) utilizam um modelo mecanístico cuja estrutura e parâmetros são obtidos a partir do código FORTRAN fornecido por Downs e Vogel (1993). São utilizados quatro controladores SISO PI para estabilizar o processo. Numa primeira abordagem são obtidos três modelos lineares, invariantes com o tempo, por diferenciação numérica do modelo nas condições de operação correspondentes aos modos de operação 1, 2 e 3. Estes modelos são depois utilizados num algoritmo de controlo predictivo linear, que inclui uma condição de troca de modelo. Esta condição quantifica o desvio observado entre as previsões do modelo e as observações, pretendendo-se utilizar o modelo que melhor represente o processo em cada uma das condições transientes. A segunda abordagem consiste na utilização de um modelo de estados partilhados (*state shared model*). Neste caso a estrutura do modelo é fixa, mas os parâmetros vão sendo adaptados consoante as condições de operação e utilizando os modelos lineares correspondentes aos três modos de operação considerados. Os resultados conseguidos com ambas as abordagens são satisfatórios, sendo os da segunda abordagem melhores no caso da transição entre os diferentes modos de operação.

Ricker (1995) determina as condições de operação óptimas para os diferentes modos de operação propostos por Downs e Vogel (1993). Estas condições são obtidas pela resolução de um problema de optimização cuja função objectivo é a função de custo fornecida por Downs e Vogel (1993) e cujas restrições são as equações do modelo em estado estacionário, bem como os limites impostos às variáveis de saída e de entrada. São ainda incluídas restrições que garantem que o caudal e a composição dos produtos correspondem aos valores desejados. O modelo utilizado é a rotina FORTRAN que simula o processo, sendo excluído o ruído das variáveis medidas. Os resultados obtidos constituem o limite inferior para este problema, uma vez que o modelo e o processo coincidem perfeitamente.

A análise das soluções obtidas permite retirar algumas conclusões, como o facto de a pressão e o nível do reactor serem as variáveis operatórias mais críticas. É também verificado que a maior parte das variáveis do processo são fortemente dependentes do

caudal e da composição do produto desejada, sendo a capacidade de produção sempre limitada pelo caudal de uma das correntes de alimentação do reactor. Daqui resulta que tentar manter qualquer uma das variáveis com um valor constante pode implicar um aumento dos custos de operação ou a limitação da capacidade de produção, pelo que deve ser evitada a utilização arbitrária de controladores SISO, e a necessidade de assegurar o controlo da composição do produto quando um dos caudais de alimentação está no seu limite superior. Os custos de operação são dominados pelas perdas na purga, pelo que estas devem ser evitadas, ou seja, a percentagem de inertes e subprodutos nesta corrente deve ser a maior possível. Ainda que toda a informação do processo seja fornecida ao algoritmo de optimização, o problema é difícil de resolver em algumas situações. Esta dificuldade reflecte-se nas próprias soluções apresentadas pelo autor. Ainda que cada uma delas corresponda a um estado estacionário, alguns dos termos diferenciais têm magnitudes da ordem de  $10^{-2}$  nas soluções apresentadas. A experiência com o modelo permite concluir que se estes estados estacionários forem utilizados como condições iniciais para a integração do modelo, rapidamente são atingidos os limites de segurança que levam a uma paragem do processo.

A abordagem ao problema da determinação das condições óptimas, do ponto de vista económico, é também considerado por [Blanco et al. \(2004\)](#), tendo também em consideração aspectos relacionados com a operabilidade do processo. O objectivo dos autores é, para o modo base, a determinação das condições óptimas de operação que correspondam a um estado estacionário estável. Para tal, ao problema de minimização do custo é adicionada a restrição de a parte real de todos os valores próprios do Jacobiano do modelo ser negativa. É utilizado o modelo simplificado proposto por [Ricker e Lee \(1995b\)](#). O custo de operação nas condições assim obtidas é cerca de dezoito vezes superior ao custo de operação mínimo, o que se deve, em grande parte, ao elevado caudal da corrente de reciclo. É também de notar que, nas condições obtidas, o caudal de produto é muito inferior ao das condições normais, o que resulta dos baixos caudais de todas as alimentações, sendo nulo o caudal de alimentação de A. Ainda que as condições de operação obtidas correspondam a um estado estacionário estável, o controlo do processo nestas condições seria particularmente difícil, uma vez que alterações nos caudais das alimentações teriam muito pouca influência sobre o processo.

[Kavouras et al. \(2005\)](#) apresentam uma metodologia para a obtenção de estados estacionários utilizando modelos dinâmicos do tipo caixa negra, isto é, sem acesso às equações que constituem o modelo. A metodologia apresentada pode ser aplicada a processos instáveis, sendo utilizado o processo TE como exemplo. No caso deste processo, o problema é resolvido recorrendo a um modelo do processo a que são acrescentadas duas equações que fixam os níveis do separador e do *stripper*, identificadas como variáveis integrantes. A adição destas duas condições leva à perda de dois graus de liberdade. Deixa de ser possível determinar simultaneamente o valor de todas as variáveis manipuláveis que resultam num estado estacionário, tendo que ser

duas destas calculadas a partir das equações do modelo.

[Duvall e Riggs \(2000\)](#) usam a estrutura de controlo descentralizado de [Ricker \(1996\)](#), ajustando os *set-points* dos controladores SISO de acordo com a solução de um problema de optimização em estado estacionário. O modelo utilizado é um modelo em estado estacionário, cujos parâmetros são ajustados de acordo com a informação do processo. A utilização deste algoritmo permite reduzir os custos de operação, quer relativamente à utilização dos *set-points* correspondentes aos modos de operação base, que relativamente à utilização dos *set-points* óptimos calculados fora de linha.

[Golshan et al. \(2005\)](#) aplicam um algoritmo de optimização em tempo real ao processo. O resultado do algoritmo é um conjunto de *set-points* que são utilizados pela estrutura de controlo descentralizada proposta por [McAvoy e Ye \(1994\)](#). O modelo utilizado, quer para a estimação dos estados (utilizando um filtro de Kalman generalizado) quer no algoritmo de optimização, é o proposto por [Ricker e Lee \(1995b\)](#). Os resultados obtidos são favoravelmente comparados com os obtidos, utilizando uma estratégia semelhante, por [Duvall e Riggs \(2000\)](#).

[Juricek et al. \(2001\)](#) comparam os modelos dinâmicos obtidos por identificação utilizando os métodos CVA, N4SID e ARX. Os autores concluem que, apesar do processo ser não-linear, as previsões conseguidas pelos modelos lineares são razoavelmente precisas para a maior parte das sete variáveis de entrada e dez variáveis de saída consideradas. Tal poder-se-à dever ao facto de ser utilizada a estrutura de controlo descentralizada proposta por [McAvoy e Ye \(1994\)](#), o que, tal como observado por [Srinivas e Arkun \(1997\)](#), poderá resultar na atenuação da não-linearidade do processo.

[Kraus et al. \(2006\)](#) abordam o problema da estimativa de estados e parâmetros utilizando um algoritmo de horizontes móveis. Nestes algoritmos a estimativa dos parâmetros e dos estados num determinado instante é obtida por resolução de um problema de optimização cuja solução minimiza a diferença entre os valores observados e os valores previstos por um modelo num intervalo de tempo finito anterior ao instante actual. Os resultados apresentados foram obtidos utilizando o modelo do processo proposto por [Jockenhövel et al. \(2003\)](#). O problema de optimização é resolvido utilizando uma estratégia de *multiple shooting* e iterações em tempo real, adaptando uma ideia semelhante proposta por [Diehl et al. \(2005\)](#) para problemas de controlo predictivo. Em cada instante de amostragem, a solução é obtida por uma iteração do método de Gauss-Newton. Desta forma, os cálculos necessários podem ser divididos numa fase de preparação, em que não é necessária a informação mais recente, e numa fase de solução, em que é utilizada a informação relativa ao instante de amostragem actual. Os resultados obtidos são favoravelmente comparados com os obtidos para o mesmo problema utilizando um filtro de Kalman generalizado, quer em termos dos erros de estimação quer em termos da eficiência computacional.

Da generalidade dos resultados apresentados podemos concluir que o processo TE,

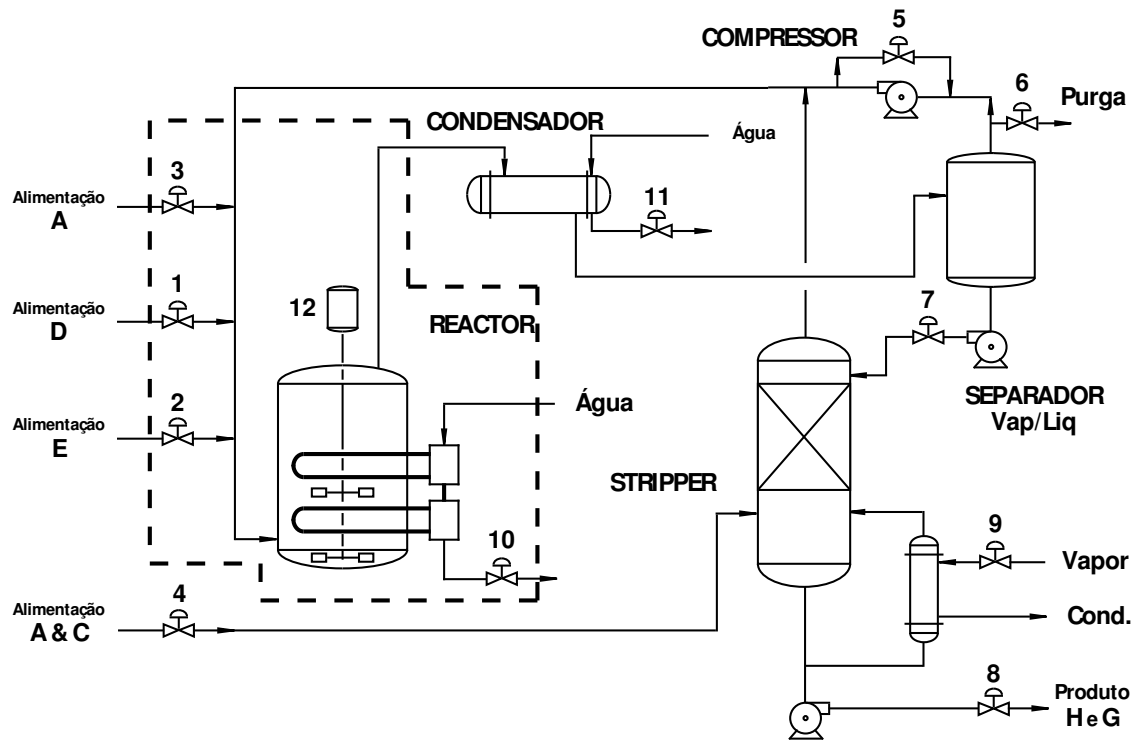


Figura 3.2 Zona do processo TE descrita pelo modelo do reator.

bem como seu modelo, possuem características que tornam o seu controlo e optimização tarefas não triviais. Entre estas, as principais são a forte interacção entre as variáveis, a não linearidade do modelo e a instabilidade do processo.

### 3.3 O reator do processo TE

As evidências obtidas, quer através de testes com o modelo, quer a partir dos resultados publicados, indicam ser o reator a unidade do processo a que se deve parte importante da instabilidade. Decidiu-se, por isso, analisar de forma mais detalhada o seu comportamento.

#### 3.3.1 Modelo do reator

De forma a ser possível estudar o comportamento do reator, foi necessária a obtenção de um modelo dinâmico isolado desta unidade. Este modelo foi obtido a partir do modelo original do processo, isolando as partes relevantes. O modelo obtido descreve a zona limitada a tracejado na Figura 3.2.

As equações diferenciais do modelo obtido são os balanços mássicos parciais ao reator e à zona de mistura dos caudais das alimentações, balanços energéticos ao reator,

à zona de mistura das correntes, ao sistema de arrefecimento, e as equações que descrevem a dinâmica das válvulas, num total de vinte e quatro. São consideradas vinte e quatro variáveis de estado, oito variáveis medidas e cinco variáveis manipuláveis. Estas variáveis, bem como o seu valor no modo base de operação, são apresentadas na Tabela 3.7. É ainda de notar que quer o caudal quer a composição da corrente de reciclo são considerados constantes e iguais aos valores correspondentes ao modo base de operação do processo.

Tabela 3.7: Variáveis do modelo do reactor.

Número	Variável	Valor estacionário	Unidades
<b>Estado</b>			
$x_1$	Moles de A na fase gasosa	10.4049	kmol
$x_2$	Moles de B na fase gasosa	4.3640	kmol
$x_3$	Moles de C na fase gasosa	7.5701	kmol
$x_4$	Moles de D na fase líquida	0.4230	kmol
$x_5$	Moles de E na fase líquida	24.1551	kmol
$x_6$	Moles de F na fase líquida	2.9426	kmol
$x_7$	Moles de G na fase líquida	154.3771	kmol
$x_8$	Moles de H na fase líquida	159.1866	kmol
$x_9$	Entalpia total do reactor	2.8085	kJ
$x_{28}$	Moles de A na zona de mistura	107.756	kmol
$x_{29}$	Moles de B na zona de mistura	29.773	kmol
$x_{30}$	Moles de C na zona de mistura	88.325	kmol
$x_{31}$	Moles de D na zona de mistura	23.039	kmol
$x_{32}$	Moles de E na zona de mistura	62.858	kmol
$x_{33}$	Moles de F na zona de mistura	5.546	kmol
$x_{34}$	Moles de G na zona de mistura	11.922	kmol
$x_{35}$	moles de H na zona de mistura	5.555	kmol
$x_{36}$	Entalpia total da zona de mistura	0.922	kmol
$x_{37}$	Temperatura da água de arref.	94.5993	°C
$x_{39}$	Posição da válvula 1	63.053	%
$x_{40}$	Posição da válvula 2	53.980	%
$x_{41}$	Posição da válvula 3	24.644	%
$x_{48}$	Posição da válvula 10	41.100	%
$x_{50}$	Velocidade de agitação do reactor	50.000	%
<b>Manipuláveis</b>			
$u_1$	Caudal da alimentação de D	63.053	%
$u_2$	Caudal da alimentação de E	53.980	%
$u_3$	Caudal da alimentação de A	24.644	%

*continua na página seguinte*

continuação da página anterior

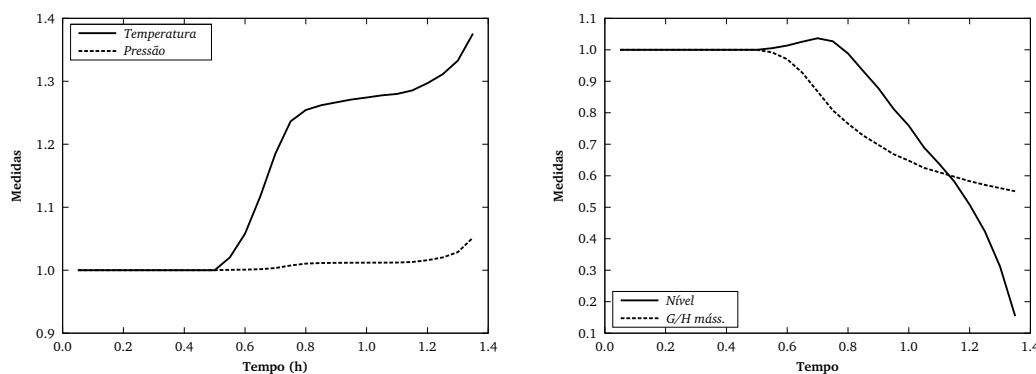
Número	Variável	Valor estacionário	Unidades
$u_{10}$	Caudal da água de arrefecimento	41.100	%
$u_{12}$	Velocidade de agitação do reactor	50.000	%
<b>Medidas</b>			
$y_1$	Caudal da alimentação de A	0.251	kscmh <sup>a</sup>
$y_2$	Caudal da alimentação de D	3664.0	kg/h
$y_3$	Caudal da alimentação de E	4509.3	kg/h
$y_6$	Caudal de alimentação do reactor	42.339	kscmh <sup>a</sup>
$y_7$	Pressão do reactor	2705.0	kPa rel.
$y_8$	Nível do reactor	75.000	%
$y_9$	Temperatura do reactor	120.40	°C
$y_{21}$	Temp. de saída da água de arref.	94.599	°C

<sup>a</sup> Thousand standard cubic meters per hour.

### 3.3.2 Resposta em ciclo aberto

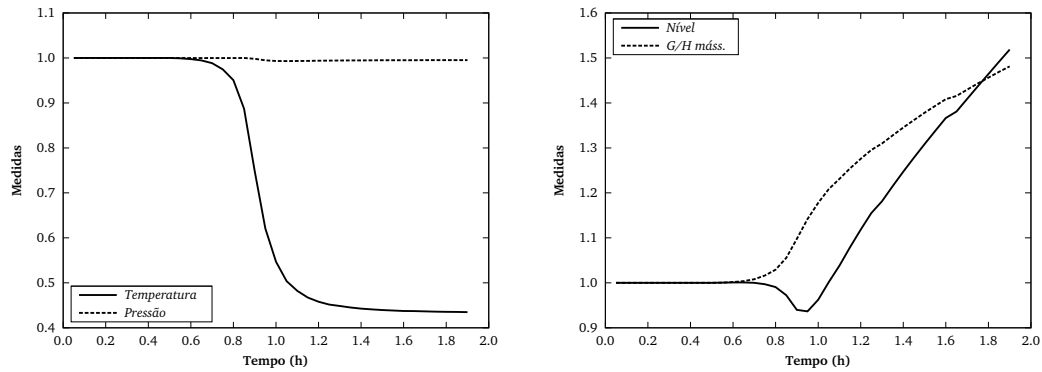
Para estudar o comportamento em ciclo aberto do reactor foram utilizadas algumas das perturbações propostas por Downs e Vogel (1993). Verificou-se que, após qualquer perturbação, não é possível atingir um novo estado estacionário sem que sejam ultrapassados os limites de segurança.

Os resultados obtidos para uma perturbação de +5°C na temperatura de entrada da água de arrefecimento e para um corte na alimentação de A são apresentados nas Figuras 3.3 e 3.4, respectivamente.



**Figura 3.3** Resposta do reactor a um aumento de +5°C na temperatura de entrada da água de arrefecimento, no instante  $t = 0.5$  h.

Os perfis apresentados são representativos dos resultados obtidos para todas as perturbações testadas. Destes resultados é possível concluir que as variáveis mais críticas



**Figura 3.4** Resposta do reactor a um corte da alimentação de A, no instante  $t = 0.5h$ .

são a pressão e o nível, sendo estas as que mais frequentemente atingem os limites de segurança.

Um aspecto relevante do comportamento do reactor é o facto de o nível apresentar resposta inversa, como pode ser observado nas Figuras 3.3 e 3.4. Tal deve-se aos fenómenos de reacção e de vaporização, cujas variações com a temperatura têm efeitos contrários. Quando a temperatura aumenta, verifica-se um aumento da velocidade das reacções, o que provoca um aumento do nível, uma vez que todos os produtos são formados no estado líquido. Por outro lado, o aumento da temperatura provoca também um aumento da vaporização, o que contribui para uma diminuição do nível. O fenómeno que responde mais rapidamente é a reacção, mas, ao manter-se o aumento da temperatura, o efeitos da vaporização tornam-se dominantes, como pode observar-se na Figura 3.3. Para uma diminuição da temperatura o comportamento é semelhante, mas com variações negativas na velocidade de reacção e na vaporização, e variações positivas no nível, como pode ser observado na Figura 3.4.

### 3.3.3 Ganhos estacionários

Para melhor compreender as relações entre as variáveis medidas e as variáveis manipuláveis, foram estudados os seus ganhos estacionários. O ganho estacionário da variável medida  $y_j$  relativamente à variável manipulada  $u_i$ , em torno do estado estacionário  $(\bar{y}, \bar{u})$ , é definido como

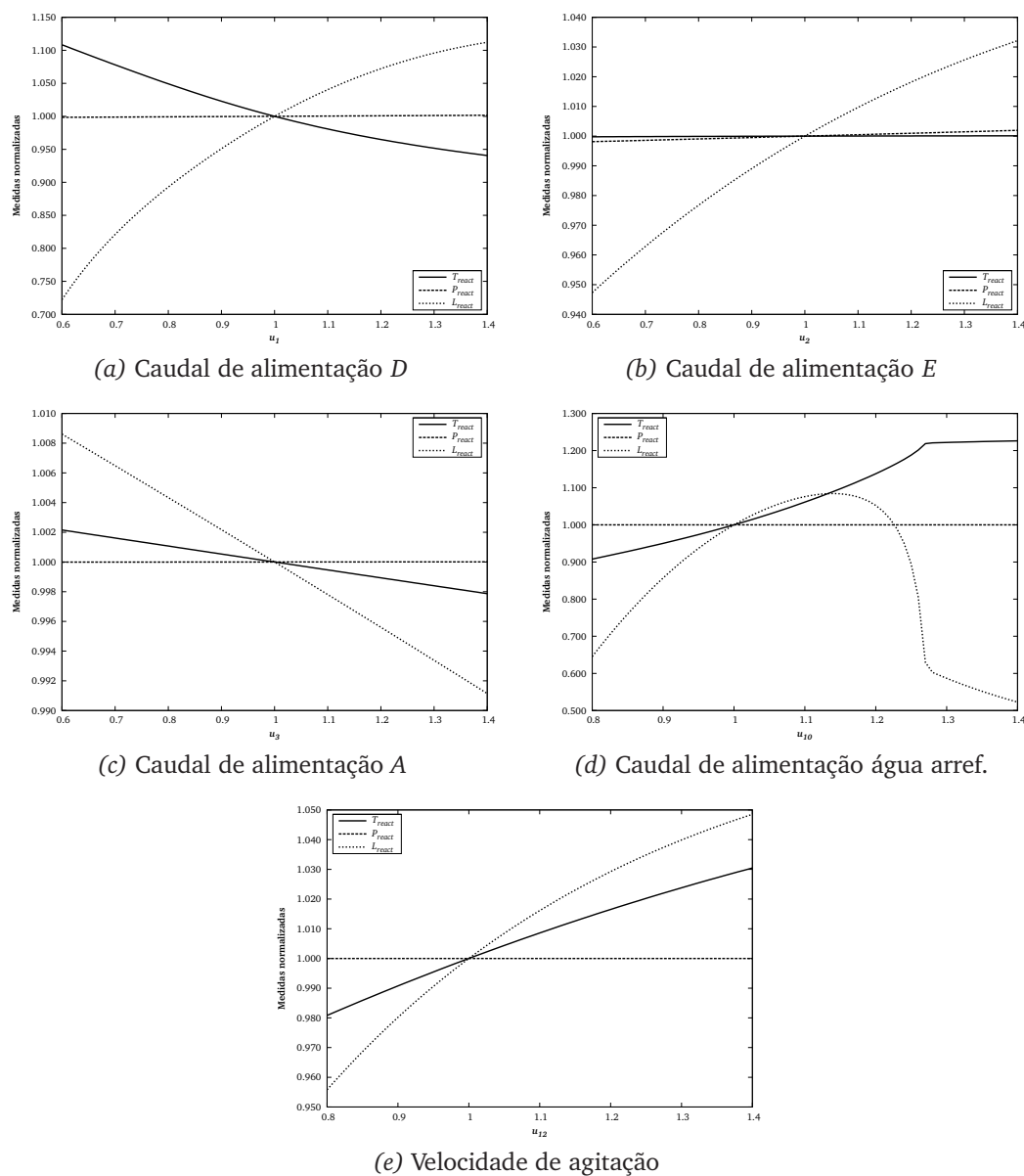
$$s_{y_j, u_i} = \lim_{\Delta u_i \rightarrow 0} \frac{y_{j,ss} - \bar{y}_j}{u_i - \bar{u}_i} = \lim_{\Delta u_i \rightarrow 0} \frac{\Delta y_{j,ss}}{\Delta u_i}$$

onde  $y_{j,ss}$  é o valor da variável medida  $j$  no estado estacionário correspondente a  $u_i$ .

Dada a forma não-analítica do modelo do processo disponível, estes ganhos foram aproximados numericamente. Para tal foram calculados novos estados estacionários para diferentes valores das variáveis manipuláveis, alterando apenas uma das variá-

veis manipuláveis e mantendo para todas as outras o seu valor no estado estacionário correspondente ao modo de operação base. Os valores das variáveis medidas nos novos estados estacionários foram representados em função dos valores da variável manipulável considerada. Os gráficos assim construídos permitem-nos concluir sobre os ganhos estacionários relativos a cada uma das variáveis manipuláveis, que serão iguais, em cada ponto, aos declives das curvas.

Na Figura 3.5 são apresentados os resultados obtidos para as cinco variáveis manipuláveis do modelo. As variáveis medidas representadas são a pressão, a temperatura e o nível do reactor. Todas as variáveis são normalizadas utilizando o seu valor no modo de operação base (Tabela 3.7).



**Figura 3.5** Valores em estado estacionário da pressão, temperatura e nível do reactor, em função das variáveis manipuláveis.



Dos resultados apresentados podemos concluir que, de forma geral, não há variações significativas dos ganhos estacionários das diferentes variáveis medidas. Exclusivamente deste ponto de vista, não seriam expectáveis grandes dificuldades à utilização de controladores lineares. Há, no entanto, algumas excepções a este comportamento. A mais notória é a do comportamento do nível com o caudal da água de arrefecimento (Figura 3.5d). Podemos verificar que o ganho estacionário é positivo até uma variação positiva de cerca de 10% no caudal da água de arrefecimento, seguindo-se uma região onde este ganho é praticamente nulo. A partir de uma variação de aproximadamente 15% do caudal de água de arrefecimento, o ganho estacionário muda de sinal, passando a ser negativo. De notar também que, nesta região, o declive da curva é muito acentuado, tendo por isso o ganho estacionário uma magnitude elevada. Ainda nesta figura podemos verificar a ocorrência de descontinuidades, quer na curva do nível quer na curva da temperatura. Estas descontinuidades devem-se à diminuição do nível do reactor que, no ponto onde ocorre a descontinuidade, deixa de ser suficiente para cobrir completamente a serpentina de arrefecimento, diminuindo a área em que ocorre transferência de calor.

É também de notar que o ganho estacionário da pressão com qualquer uma das variáveis de entrada é muito pequeno. Tal é devido à simplificação efectuada neste modelo ao considerar que o caudal, temperatura e a composição da corrente de reciclo são constantes. Uma vez que o caudal mássico do reciclo é bastante superior ao das alimentações individuais, é natural que estes tenham relativamente pouca influência no comportamento da pressão, que depende fortemente do quantidade de gás no reactor. Relativamente ao caudal da água de arrefecimento verificamos que as variações da temperatura, do nível (e logo do volume de gás), e do número de moles da fase gasosa são de forma a manter a pressão aproximadamente constante.

No gráfico das sensibilidades relativas ao caudal da água de arrefecimento, Figura 3.5d, verifica-se um aumento da temperatura em estado estacionário com o aumento daquele caudal. Este comportamento pode ser explicado pelo balanço energético ao reactor que, de forma simplificada, pode ser escrito como

$$\dot{H} = H_0 F_0 - HF + \Delta H_R - Q_c$$

onde  $H$  é a entalpia total,  $F_0$  é o caudal da alimentação,  $H_0$  a respectiva entalpia,  $F$  é o caudal de saída,  $H$  a respectiva entalpia,  $\Delta H_R$  é o termo de geração de calor devido às reacções e  $Q_c$  é a quantidade de calor removido. Com o aumento do caudal da água de arrefecimento aumenta também a quantidade de calor transferido. Para se atingir um estado estacionário é necessário que o termo  $\Delta H_R - HF$  contrabalance aquela variação, resultando num aumento da temperatura. Deve ser salientado que daqui não podemos concluir que um aumento do caudal de arrefecimento provoque um aumento da temperatura do reactor. De facto, os testes de simulação realizados demonstram que ao fazer uma perturbação positiva no caudal ocorre uma quase ime-

diata diminuição da temperatura. Devido à instabilidade do sistema, discutida na secção seguinte, o sistema não atinge nenhum estado estacionário após realização da perturbação.

Nos gráficos apresentados nas Figuras 3.5a, 3.5b e 3.5c, relativos aos caudais das alimentações de  $D$ ,  $E$  e  $A$ , podemos verificar que o ganho do nível com estes caudais apenas é negativo relativamente ao de  $A$ . Tal deve-se ao facto de o reagente existente no reactor em maior quantidade ser  $A$ , o que implica que um aumento no caudal de  $A$  provoque alterações relativamente pequenas na sua pressão parcial, quando comparadas com as alterações semelhantes efectuadas nos caudais dos outros reagentes. Assim, um aumento do caudal de  $A$  contribui sobretudo para uma diminuição das pressões parciais dos restantes reagentes o que implica uma diminuição das velocidades de reacção, uma vez que o aumento da pressão parcial de  $A$  não é significativo.

Na Figura 3.5c podemos concluir que todas as variáveis medidas são muito pouco sensíveis a variações no caudal de alimentação de  $A$ , sendo a variação máxima de 0.9% no nível para uma variação de -50% no caudal. Tal deve-se ao facto de o caudal de  $A$  ( $22.4\text{kg h}^{-1}$ ) ser cerca de cem vezes menor que os restantes caudais de alimentação do reactor.

### 3.3.4 Linearização do modelo do reactor

Para confirmar a instabilidade do reactor em ciclo aberto, o modelo foi linearizado, de forma a obter uma representação no espaço de estados com a forma

$$\dot{x}' = Ax' + Bu' \quad (3.2a)$$

$$y' = Cx' + Du' \quad (3.2b)$$

onde  $x'$ ,  $y'$  e  $u'$  representam os desvios das variáveis relativamente ao ponto em torno do qual é feita a linearização. O modelo linear do reactor foi obtido por diferenciação numérica das equações do modelo em torno do modo de operação base.

Dada a forma não analítica do modelo do reactor, recorreu-se à utilização de diferenças finitas *forward*, utilizando o algoritmo proposto por Gill et al. (1983). Este método foi escolhido por permitir boas aproximações às derivadas de primeira ordem, sem um excessivo esforço de cálculo.

Os valores próprios da matriz de estado do reactor, matriz  $A$  em (3.2a), são apresentados na Tabela 3.8. Apenas é apresentada a parte real, uma vez que todos têm parte imaginária nula.

A matriz de estado do reactor tem um valor próprio positivo, o que confirma a instabilidade em ciclo aberto do reactor, em sentido estrito. É também de notar a grande

**Tabela 3.8** Valores próprios da matriz de estado do modelo linearizado do reactor.

Parte Real	Parte Real
-1690.	-54.80
-720.0 <sup>a</sup>	-26.18
-600.0	-24.43
-530.6	-12.44 <sup>b</sup>
-450.0 <sup>a</sup>	12.20
-115.7	-11.61
-89.16	-2.354
-84.91	-0.6137

<sup>a</sup> multiplicidade 2

<sup>b</sup> multiplicidade 7

diferença de magnitudes entre os valores próprios. Esta diferença resulta de existência de variáveis com velocidades de resposta muito diferentes. É possível identificar as variáveis associadas aos valores próprios com menor parte real (maior velocidade de resposta) como sendo as cinco variáveis de estado correspondentes à dinâmica das variáveis manipuladas.

A utilização de um modelo linear do processo implica que, da sua análise, não possam ser retiradas quaisquer conclusões sobre a estabilidade de Lyapunov ou a estabilidade assintótica global. A análise dos valores próprios da matriz de estado obtida nas condições acima referidas, apenas permite concluir sobre a estabilidade assintótica local e as conclusões obtidas apenas se podem aplicar ao modelo não-linear numa pequena região em torno do modo normal de operação (Willems, 1970).

### 3.4 Caracterização do processo

Para que a supervisão de um processo químico seja feita de forma adequada é importante conhecer previamente as suas características; nesta tarefa, a existência de um modelo mecanístico detalhado, que permita simular diferentes condições de operação, é uma ferramenta extremamente valiosa. Utilizando este modelo é possível determinar as condições óptimas, quer do ponto de vista económico para diferentes regimes de operação, quer do ponto de vista da maximização da capacidade de produção. A determinação destas condições permite ainda identificar quais as restrições do processo que as condicionam, o que é importante para a selecção de estruturas de controlo adequadas, uma vez que estas restrições corresponderão a graus de liberdade que deixam de poder ser utilizados para o controlo do processo. Permite também identificar limitações específicas do processo, o que pode ser útil, por exemplo, na eliminação de estrangulamentos do processo (*debottlenecking*). Para além destes as-

pectos, para o projecto de estruturas de controlo é também útil a caracterização do comportamento dinâmico do processo, o que pode envolver, numa primeira etapa, a análise dos ganhos estacionários de todas as variáveis medidas relativamente às variáveis que podem ser manipuladas.

Com este objectivo, foram determinadas as condições de operação do processo TE que permitem maximizar a capacidade de produção, para as diferentes misturas de  $G$  e  $H$  pretendidas no produto final. Uma vez que nem sempre será desejável operar nestas condições, foram também determinadas as condições de operação que, para diferentes volumes de produção, permitem a minimização dos custos. Foram também caracterizadas as regiões de operação viáveis, de forma a poder verificar como é que as condições do processo variam nos seus extremos e no seu interior. Finalmente, foram obtidas condições de operação para diferentes valores das variáveis manipuláveis, de forma a analisar os ganhos estacionários das variáveis medidas relativamente a cada uma daquelas variáveis. Os resultados obtidos são apresentados nas Secções seguintes.

### 3.4.1 Linearização do modelo do processo

À semelhança do que foi feito para o modelo do reactor (Secção 3.3.4), também o modelo do processo foi linearizado, com o objectivo de obter a matriz de estado e calcular os respectivos valores próprios para avaliar a estabilidade do processo em ciclo aberto e a existência de modos integrantes. Esta análise é importante para a restante caracterização do processo, uma vez que a existência de modos integrantes implica a singularidade da matriz de estado. Sendo esta matriz singular, os métodos de resolução de sistema de equações não-lineares ou de problemas de optimização em que esta matriz surja como um conjunto de restrições terão que ser adaptados a esta situação ou, de alguma forma, reformulados.

A matriz de estado do processo foi obtida utilizando diferenças finitas *forward* e o algoritmo proposto por Gill et al. (1983). Os valores próprios da matriz de estado assim obtida são apresentados na Tabela 3.9.

Podemos verificar que, tal como o reactor, o processo tem pólos com parte real positiva (quatro pólos, correspondentes a dois pares de valores próprios complexos). Existe também um valor próprio que pode ser considerado nulo ( $2.1 \times 10^{-8}$ ). Há ainda um segundo valor próprio que poderá ser considerado nulo ( $-8.3 \times 10^{-4}$ ), uma vez que a sua magnitude é semelhante à do limite do erro resultante do algoritmo aplicado para o cálculo da matriz de estado. Conclui-se que o processo é instável, no sentido estrito, em ciclo aberto e que possui dois modos integrantes. Devem aqui ser aplicadas as mesmas reservas referidas na interpretação dos mesmos resultados para o reactor (Secção 3.3.4).

**Tabela 3.9** Valores próprios da matriz de estado do modelo linearizado do processo.

Parte		Parte	
Real	Imaginária	Real	Imaginária
-1968.07		-26.228	
-726.369		-13.116	
-720. <sup>a</sup>		-12.833	
-600.		-11.619	-8.49
-529.699		-11.619	8.49
-514.286		-10.760	
-450. <sup>b</sup>		-9.935	
-400.		-8.921	
-139.164		-5.178 <sup>c</sup>	
-114.194		3.065	-5.084
-87.025		3.065	5.084
-81.858		-1.222	
-37.220	-17.119	$-1.153 \times 10^{-1}$	
-37.220	17.119	$-8.134 \times 10^{-2}$	
-34.797	-14.934	$-7.452 \times 10^{-2}$	
-34.797	14.934	$-8.338 \times 10^{-4}$	
-30.		$2.103 \times 10^{-8}$	
-29.275		$2.545 \times 10^{-2}$	$-1.55 \times 10^{-1}$
-28.835		$2.545 \times 10^{-2}$	$1.55 \times 10^{-1}$

<sup>a</sup> multiplicidade 6<sup>b</sup> multiplicidade 2<sup>c</sup> multiplicidade 7

É de notar que os valores próprios obtidos não coincidem totalmente com os resultados apresentados por outros autores que, por sua vez, não são coincidentes entre si. [McAvoy \(1998\)](#) refere a existência de dois pares de valores próprios complexos com parte real positiva. [Ricker e Lee \(1995b\)](#) referem a existência de valores próprios com valores entre -1968 a 3.07, enquanto que [Banerjee e Arkun \(1995\)](#) referem a existência de seis valores próprios com parte real positiva. A nossa experiência indica que os valores próprios são muito sensíveis às aproximações feitas na linearização do modelo, podendo ser essa a razão para as diferenças referidas. Há, no entanto, algo comum a todos estes resultados que é a existência de pelo menos um valor próprio com parte real positiva.

### 3.4.2 Condições óptimas de operação

Dos seis modos de operação apresentados na Tabela 3.1, três correspondem ao caudal máximo de produto. Para caracterizar estes modos de operação é necessário deter-

minar as condições de operação que permitem a obtenção destes caudais máximos, sendo para tal necessária a resolução de um problema de optimização. Neste problema de optimização, a função objectivo será o caudal de produto e as restrições serão as equações do modelo, com os termos diferenciais iguais a zero. Será ainda necessário considerar uma restrição que garanta a razão pretendida entre os caudais mássicos de  $G$  e  $H$  no produto.

Para além de estabelecerem objectivos de controlo, [Downs e Vogel \(1993\)](#) estabelecem ainda um critério económico que pode ser utilizado na optimização do processo. Este critério é o custo total de operação, sobretudo devido à perda de produtos e reagentes, que ocorre na purga. Os reagentes condensáveis ( $D$  e  $E$ ) também são perdidos no produto. É também atribuído um custo à quantidade de  $F$  produzido, contabilizando o custo dos reagentes gastos nas reacções secundárias. Também são considerados os custos do vapor utilizado no *stripper* e o trabalho realizado pelo compressor. O custo total de operação pode ser calculado pela expressão ([Downs e Vogel, 1993](#))

$$C_{\text{total}} = 0.0536y_{20} + 0.0318y_{19} + 0.47791y_{10} [2.209y_{29} + 6.177y_{31} + 22.06y_{32} + 14.56y_{33} + 17.89y_{34} + 30.44y_{35} + 22.94y_{36}] + 4.541x_{46} [0.2206y_{37} + 0.1456y_{38} + 0.1789y_{39}]$$

onde  $4.541x_{46}$  é o caudal de produto, em kmol/h, e  $C_{\text{total}}$  em \$/h. Com esta função objectivo é possível obter as condições de operação que, para os modos de operação 1, 2 e 3 da Tabela [3.1](#), minimizam os custos.

Na resolução dos problemas de optimização foram, inicialmente, utilizadas as rotinas originais do modelo, em FORTRAN, não tendo sido possível a obtenção de soluções com qualquer um dos algoritmos testados. Esta falha deveu-se ao facto de o modelo ter sido utilizado como um modelo do tipo *caixa negra*, não tendo sido consideradas algumas especificidades deste processo importantes no contexto da aplicação de métodos de optimização estacionária ou de solução de sistemas de equações não-lineares. Estas questões são referidas de seguida, na descrição do modelo reformulado que foi utilizado. A própria forma como o modelo é implementado nas rotinas em FORTRAN dificulta esta tarefa, uma vez que não é facilmente possível o acesso a toda a informação do modelo.

Optou-se, por isso, pela utilização do sistema GAMS ([Rosenthal, 2010](#)), o que implicou a reescrita das equações do modelo. Para além disso, porque as rotinas originais se destinam à simulação dinâmica do processo, foram também introduzidas algumas alterações de forma a tornar o modelo mais adequado à obtenção de condições de estado estacionário. A principal alteração introduzida foi a eliminação das variáveis de estado correspondentes às quantidades totais das fases líquidas no separador e no *stripper*. A necessidade desta alteração deve-se ao facto de ambas serem variáveis in-

tegrantes, a que se devem os dois pólos nulos da matriz de estado, não tendo qualquer influência no estado estacionário do processo. A presença destas variáveis inviabiliza a determinação de estados estacionários, uma vez que o Jacobiano do modelo será singular. Como consequência da eliminação destas variáveis, foi também necessário eliminar as quantidades parciais de cada composto nas fases líquidas destas duas unidades. Não sendo consideradas estas variáveis, foi necessária a inclusão de condições que asseguram a consistência das soluções obtidas, nomeadamente que a soma das fracções molares nas fases líquidas é igual à unidade. Esta alteração é diferente da utilizada por Ricker (1995) na solução do mesmo problema. Este autor adiciona duas restrições ao problema de optimização que fixam o nível destas unidades em 50%. Estas restrições podem ter implicações sobre a composição e a temperatura das fases líquidas, uma vez que estas variáveis influenciam a densidade e, juntamente com a quantidade total de líquido, determinam o nível.

No modelo dinâmico, as entalpias totais, em cada uma das unidades, são apenas utilizadas no balanços energéticos para quantificar a acumulação de energia, sendo depois utilizadas para calcular a temperatura, a partir da qual são calculadas as restantes propriedades necessárias à simulação. Assim, não é necessário considerar as variáveis correspondentes às entalpias no modelo estacionário, sendo necessário apenas convergir as equações nas quais intervém a temperatura. Foram também eliminadas as doze variáveis de estado relativas à dinâmica das válvulas, uma vez que, em estado estacionário, estas são iguais às variáveis manipuláveis.

Outra alteração necessária relativamente ao modelo original é a forma como é definida a variação da área de transferência de calor com o nível, no reactor. No modelo original, a fracção da área total da serpentina onde efectivamente ocorre transferência de calor é determinada pela condição

$$uar(v) = \begin{cases} 1, & \text{se } v \geq 390 \\ 0.025 \frac{v}{7.8} - 0.25, & \text{se } v < 390 \end{cases}$$

onde  $v$  é determinada pela quantidade e densidade da fase líquida do reactor. Quando  $uar(v) = 1$ , o dispositivo de transferência de calor encontra-se totalmente coberto pela fase líquida. Na linguagem GAMS não é possível a utilização deste tipo de condições discretas na resolução de problemas não-lineares contínuos (NLP), pelo que foi necessário o recurso a uma aproximação. Baseada na aproximação hiperbólica proposta por Balakrishna e Biegler (1992), a aproximação utilizada foi

$$\phi(v) = 1 - 0.5\sqrt{f(v)^2 + \epsilon^2} + 0.5f(v)$$

com

$$f(v) = 1 - \left(0.025 \frac{v}{7.8} - 0.25\right)$$



e  $\epsilon = 0.001$ . O erro relativo máximo desta aproximação é de aproximadamente 0,05%.

Um das dificuldades verificadas na obtenção de resultados deve-se ao mau *scaling* do modelo original. Este problema foi atenuado pela introdução de factores de escala, para variáveis e equações, e pela alteração das unidades de algumas das variáveis. Para possibilitar a comparação com os resultados obtidos por Ricker (1995), foram considerados os limites utilizados por este autor para as variáveis do processo que estão sujeitas a restrições, e não os apresentados na Tabela 3.6. Todos os resultados apresentados nesta secção foram obtidos estabelecendo o limite inferior do nível do reactor em 65% e o limite superior para a pressão do reactor em 2800 kPa. Para todas as variáveis manipuláveis os limites inferior e superior são 1% e 100%, respectivamente.

Na Tabela 3.10 é apresentado o resumo dos resultados obtidos para a maximização do caudal de produto. Os valores em itálico correspondem a variáveis que se encontram no limite inferior ou superior.

Os resultados obtidos mostram que o caudal máximo é conseguido, em boa parte, pela saída de reagentes não convertidos na corrente de saída. Isto reflecte-se no elevado valor do custo associado às perdas no produto. Sendo os produtos de interesse  $G$  e  $H$ , será mais relevante o problema de maximização do caudal destes produtos na corrente de saída. Estando os dois caudais mássicos relacionados pela restrição que define a razão pretendida entre ambos, a função objectivo deste problema pode ser simplesmente o caudal de  $G$  na corrente de saída. Os resultados obtidos são apresentados na Tabela 3.11.

Podemos verificar que, em todas as soluções obtidas, a pressão do reactor se encontra no seu limite superior e o nível do reactor no seu limite inferior. Desta forma é possível aumentar a quantidade de reagentes no reactor e, em consequência, maximizar as taxas de reacção. Também podemos verificar que existem sempre caudais de alimentações de reagentes que se encontram no limite superior. Desta análise também podemos concluir que a produção de  $G$  é limitada pela disponibilidade de  $D$ , enquanto que a produção de  $H$  é limitada pela disponibilidade de  $E$ . Ainda que não seja limitante, é de notar que o caudal de  $C$  se encontra próximo do seu limite superior no regime  $50G/50H$ . Nos modos de operação  $50G/50H$  e  $90G/10H$ , o caudal de água de arrefecimento do separador está no seu limite superior.

Para este problema em particular, os resultados obtidos diferem significativamente dos obtidos por Ricker (1995), sendo que esta diferença se deve à forma como o problema foi resolvido. Aquele autor começa por identificar qual é o caudal de alimentação limitante para um determinado regime de operação. O caudal máximo é depois determinado resolvendo o problema de minimização do custo, com o caudal limitante constante e igual ao seu limite superior. As soluções aqui apresentadas foram



**Tabela 3.10** Condições de operação para caudal máximo de produto.

	50G/50H	10G/90H	90G/10H
<b>Custos:</b>			
Purga (\$/h)	98.31	112.41	119.46
Perdas no produto (\$/h)	465.04	56.30	288.06
Compressor (\$/h)	19.36	18.87	14.24
Vapor (\$/h)	0.33	16.67	0.28
Total (\$/h)	583.03	204.24	422.04
Unitário (¢/h)	2.44	1.37	3.31
<b>Caudal de produto (kg/h):</b>			
Total	23894.62	14864.83	12756.95
<i>G</i>	11180.49	1463.58	10762.66
<i>H</i>	11180.49	13172.24	1195.85
<b>Variáveis manipuláveis (%):</b>			
$u_1$ (caudal <i>D</i> )	<i>100.00</i>	13.09	<i>100.00</i>
$u_2$ (caudal <i>E</i> )	<i>100.00</i>	<i>100.00</i>	17.24
$u_3$ (caudal <i>A</i> )	48.74	33.79	20.73
$u_4$ (caudal <i>C</i> )	96.06	58.39	55.32
$u_5$ (valv. reciclo)	14.01	54.58	19.24
$u_6$ (valv. purga)	35.92	38.66	30.08
$u_7$ (valv. separador)	68.15	39.42	43.91
$u_8$ (valv. <i>stripper</i> )	80.19	44.88	46.22
$u_9$ (valv. vapor)	<i>1.00</i>	<i>100.00</i>	<i>1.00</i>
$u_{10}$ (arref. reactor)	80.11	28.13	41.55
$u_{11}$ (arref. condensador)	<i>100.00</i>	<i>100.00</i>	<i>100.00</i>
$u_{12}$ (agitação reactor)	35.63	41.60	<i>100.00</i>
<b>Variáveis restringidas:</b>			
$y_7$ ( $P_{\text{reactor}}$ , kPa)	<i>2800.</i>	<i>2800.</i>	<i>2800.</i>
$y_8$ ( $L_{\text{reactor}}$ , %)	<i>65.00</i>	<i>65.00</i>	<i>65.00</i>
$y_9$ ( $T_{\text{reactor}}$ , °C)	122.92	125.06	111.27
$y_{12}$ ( $L_{\text{separador}}$ , %)	46.91	53.09	46.77
$y_{15}$ ( $L_{\text{stripper}}$ , %)	44.73	57.14	40.12

*Nota: os valores apresentados em itálico são variáveis que se encontram num dos seus limites.*

obtidas por maximização directa do caudal de produto.

Comparando as Tabelas 3.10 e 3.11, para a mistura 50G/50H podemos verificar que a diferença no caudal de *G* é relativamente pequena (aproximadamente 26 kg/h). No entanto, a diferença de custo é bastante significativa (cerca de 300\$/h). Isto indica que o custo total de operação, nesta região é fortemente dependente das condições de operação e do caudal de produto pretendido.

Utilizando o custo total de operação como função objectivo, foram obtidas as condições de operação que correspondem ao custo mínimo de operação para os modos de operação 1, 2 e 3 da Tabela 3.1. Os resultados obtidos são apresentados na Tabela 3.12.

Em todas as soluções correspondentes ao custo mínimo de operação, quer a válvula

**Tabela 3.11** Condições de operação para caudal máximo de G no produto.

	50G/50H	10G/90H	90G/10H
<b>Custos:</b>			
Purga (\$/h)	100.53	138.80	23.70
Perdas no produto (\$/h)	137.45	39.47	7.93
Compressor (\$/h)	19.63	19.38	16.26
Vapor (\$/h)	22.45	14.69	12.59
Total (\$/h)	280.07	212.34	60.49
Unitário (¢/h)	1.22	1.43	0.48
<b>Caudal de produto (kg/h):</b>			
Total	22925.11	14836.94	12537.29
G	11206.25	1465.90	11211.42
H	11206.25	13193.07	1245.71
<b>Variáveis manipuláveis (%):</b>			
$u_1$ (caudal D)	<i>100.00</i>	13.14	<i>100.00</i>
$u_2$ (caudal E)	88.01	<i>100.00</i>	9.39
$u_3$ (caudal A)	47.82	36.17	21.38
$u_4$ (caudal C)	96.49	58.83	57.68
$u_5$ (valv. reciclo)	18.15	52.80	69.72
$u_6$ (valv. purga)	37.23	46.63	10.74
$u_7$ (valv. separador)	65.52	37.98	33.59
$u_8$ (valv. stripper)	75.43	44.71	44.24
$u_9$ (valv. vapor)	<i>100.00</i>	<i>100.00</i>	<i>100.00</i>
$u_{10}$ (arref. reactor)	74.77	28.58	42.44
$u_{11}$ (arref. condensador)	<i>100.00</i>	41.61	<i>100.00</i>
$u_{12}$ (agitação reactor)	40.10	41.60	<i>100.00</i>
<b>Variáveis restringidas:</b>			
$y_7$ ( $P_{\text{reactor}}$ , kPa)	<i>2800.</i>	<i>2800.</i>	<i>2800.</i>
$y_8$ ( $L_{\text{reactor}}$ , %)	<i>65.00</i>	<i>65.00</i>	<i>65.00</i>
$y_9$ ( $T_{\text{reactor}}$ , °C)	127.37	125.86	120.12
$y_{12}$ ( $L_{\text{separador}}$ , %)	47.61	53.80	43.79
$y_{15}$ ( $L_{\text{stripper}}$ , %)	47.81	57.93	43.99

Nota: os valores apresentados em *itálico* são variáveis que se encontram num dos seus limites.

de reciclo do compressor quer a válvula de vapor do *stripper* se encontram no seu limite inferior. Desta forma reduzem-se os custos associados ao trabalho realizado pelo compressor e ao consumo de vapor. Novamente, a pressão e o nível do reactor se encontram nos limites superior e inferior, respectivamente, em toda as condições de operação com custo mínimo.

Todos os problemas foram resolvidos utilizando o algoritmo CONOPT (Drud, 1994). Para tentar assegurar que os resultados obtidos não correspondem a óptimos locais, foi utilizado o algoritmo OQNLP (Ugray et al., 2005), que implementa uma estratégia de pesquisa que consiste em resolver o problema partindo de diferentes condições iniciais. Em nenhum dos problemas foi detectada a existência de soluções óptimas significativamente diferentes. É, no entanto, de salientar o facto de existirem diferentes valores das variáveis  $u_{10}$  (água de arrefecimento do reactor) e  $u_{12}$  (velocidade de

**Tabela 3.12** Condições de operação para custo mínimo.

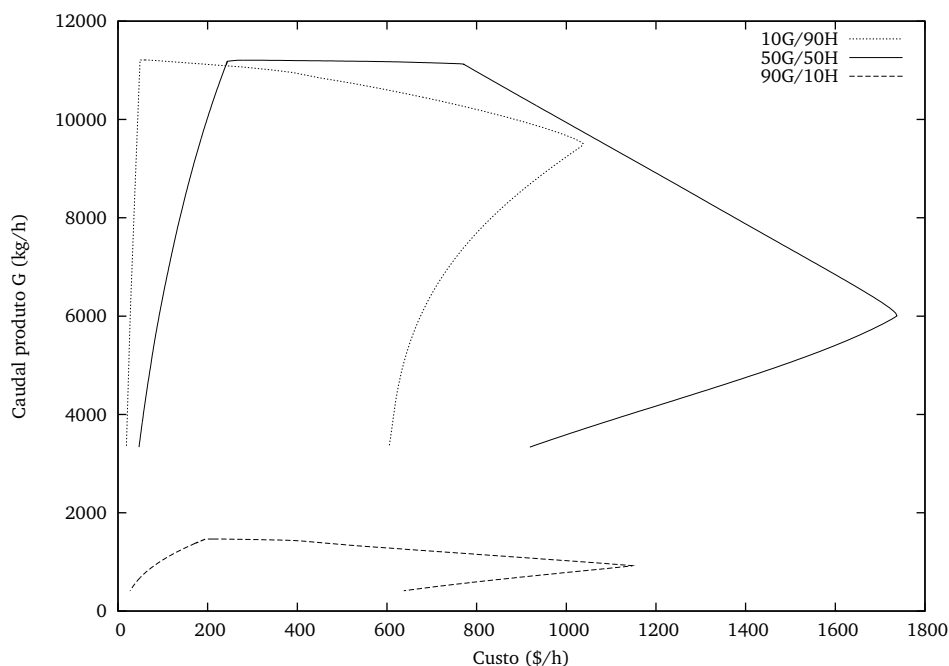
	50G/50H	10G/90H	90G/10H
<b>Custos:</b>			
Purga (\$/h)	73.87	130.33	29.07
Perdas no produto (\$/h)	25.52	36.29	6.47
Compressor (\$/h)	14.95	14.72	14.90
Vapor (\$/h)	0.15	0.16	0.12
Total (\$/h)	114.49	181.50	50.56
Unitário (¢/h)	0.81	1.27	0.45
<b>Caudal de produto (kg/h):</b>			
Total	14213.69	14243.40	11180.88
<i>G</i>	7038.00	1408.00	10000.00
<i>H</i>	7038.00	12669.00	1111.00
<b>Variáveis manipuláveis (%):</b>			
$u_1$ (caudal <i>D</i> )	63.00	12.65	89.35
$u_2$ (caudal <i>E</i> )	53.21	96.33	8.41
$u_3$ (caudal <i>A</i> )	26.28	30.46	18.50
$u_4$ (caudal <i>C</i> )	60.63	56.16	51.47
$u_5$ (valv. reciclo)	1.00	1.00	1.00
$u_6$ (valv. purga)	25.81	44.44	11.32
$u_7$ (valv. separador)	37.31	35.84	29.92
$u_8$ (valv. <i>stripper</i> )	46.49	42.91	39.45
$u_9$ (valv. vapor)	1.00	1.00	1.00
$u_{10}$ (arref. reactor)	42.46	26.79	39.56
$u_{11}$ (arref. condensador)	12.46	12.94	7.85
$u_{12}$ (agitação reactor)	30.44	48.62	47.34
<b>Variáveis restringidas:</b>			
$y_7$ ( $P_{\text{reactor}}$ , kPa)	2800.	2800.	2800.
$y_8$ ( $L_{\text{reactor}}$ , %)	65.00	65.00	65.00
$y_9$ ( $T_{\text{reactor}}$ , °C)	122.88	124.22	121.29
$y_{12}$ ( $L_{\text{separador}}$ , %)	51.62	56.72	46.92
$y_{12}$ ( $L_{\text{stripper}}$ , %)	49.93	57.53	44.11

*Nota: os valores apresentados em itálico são variáveis que se encontram num dos seus limites.*

agitação do reactor) que conduzem à mesma solução. Isto deve-se ao facto de ambas as variáveis influenciarem a quantidade de calor transferido no reactor, sendo possível que esta quantidade de calor assuma o mesmo valor para diferentes combinações daquelas duas variáveis.

### 3.4.3 Regiões de operação viáveis

Os resultados obtidos na secção anterior demonstram que, para além das condições de operação óptimas, é também importante caracterizar o comportamento do processo em torno destas. Para isso foram determinadas as regiões de operação viáveis para cada uma das três razões  $G/H$ . Estas regiões foram determinadas fazendo va-

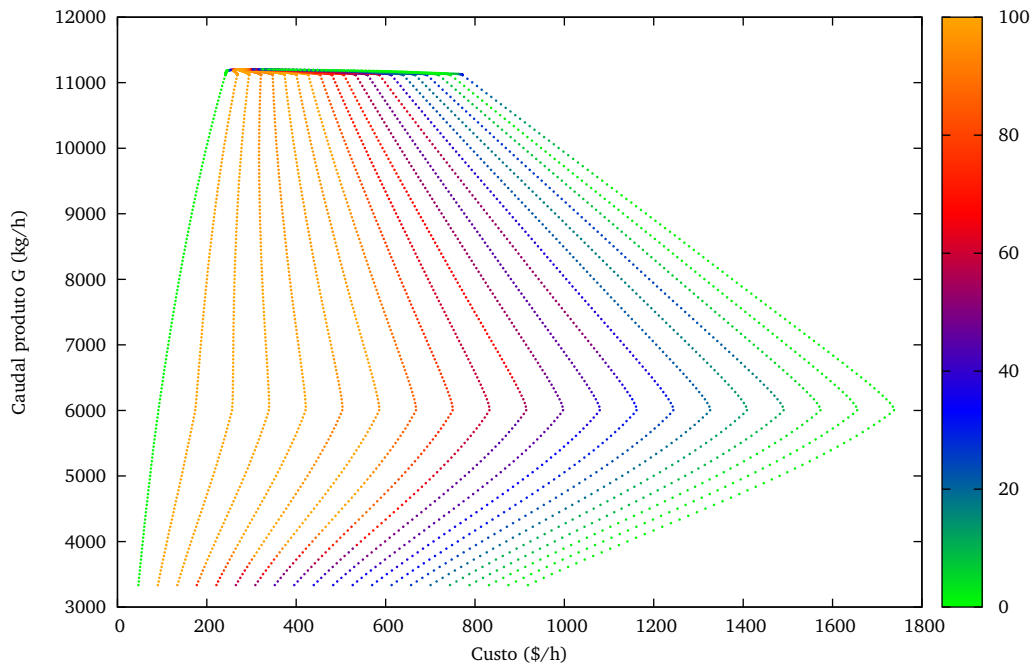


**Figura 3.6** Regiões de operação viáveis para as três razões  $G/H$  no produto.

riar, a partir do máximo, o caudal mássico de  $G$  e, com este valor fixo, resolvendo os problemas de maximização e minimização do custo total. Todos os problemas foram resolvidos utilizando o algoritmo OQNLP, de forma a obter a solução global. Foram também identificados alguns óptimos locais, sobretudo na solução do problema de maximização do custo. Nos problemas de minimização não foram encontradas soluções significativamente diferentes. As regiões de operação obtidas são apresentadas na Figura 3.6, para as três razões  $G/H$ .

Como já indiciado pelos resultados apresentados na secção anterior, verifica-se que, para todos os regimes de operação, a região em torno do caudal máximo é relativamente plana. Ou seja, com uma ligeira diminuição do caudal de  $G$  é possível uma redução significativa dos custos de operação. Em todos os pontos foram identificados diferentes valores das variáveis  $u_{10}$  e  $u_{12}$  que resultam na mesma solução, o que se deve às razões já apresentadas.

Para verificar a viabilidade das regiões de operação apresentadas, foram determinadas as condições de operação para um conjunto de pontos no seu interior. As condições de operação nestes pontos foram obtidas fixando o caudal mássico de  $G$  e estabelecendo o custo pretendido. Os valores do caudal e do custo foram estabelecidos de forma a varrer todo o interior das regiões apresentadas na Figura 3.6. Para cada um dos



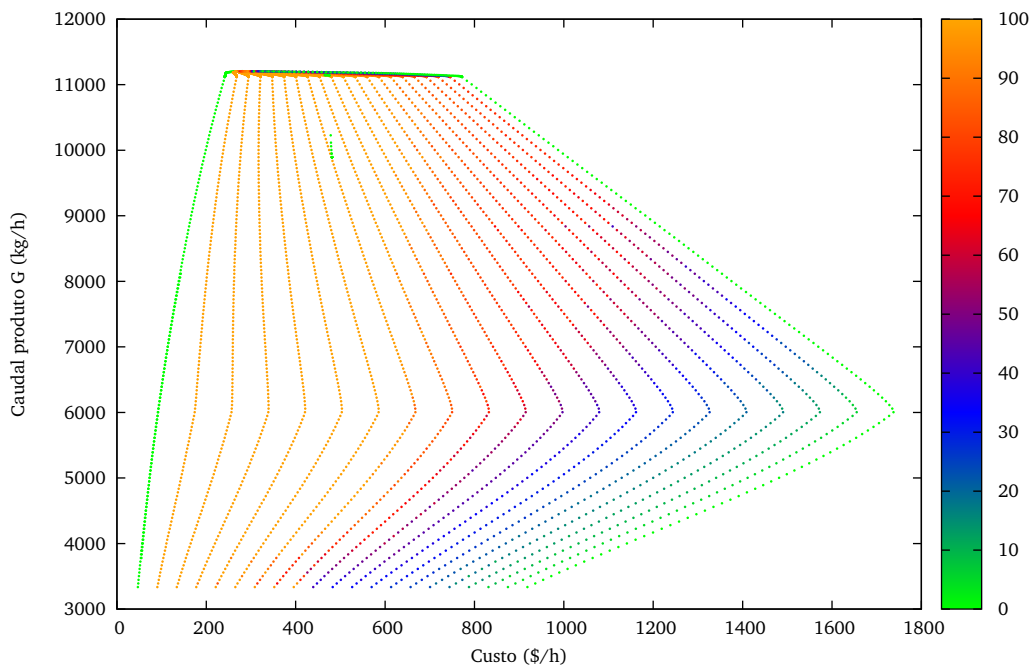
**Figura 3.7** Posição da válvula de vapor do *stripper* ( $u_9$ ) em função do custo total e do caudal mássico de  $G$ , no regime de operação  $50G/50H$ .

pontos foi resolvido o problema

$$\begin{aligned} \min \quad & (C_{\text{total}} - C_{\text{ref}})^2 \\ \text{s.a} \quad & \text{equações do modelo;} \\ & \text{caudal } G \text{ pretendido;} \\ & \text{composição pretendida,} \end{aligned}$$

em que  $C_{\text{ref}}$  é o custo pretendido para a solução. Para cada valor do caudal de  $G$ , o problema foi resolvido para dezanove valores diferentes de  $C_{\text{ref}}$ , igualmente espaçados entre os custos mínimo e máximo. A título de exemplo, é apresentada, na Figura 3.7, a variação da variável manipulável  $u_9$  (válvula de vapor do *stripper*), no regime de operação  $50G/50H$ . Os mesmos resultados, para as restantes variáveis manipuláveis, são apresentados no Apêndice A.

Na resolução deste conjunto de problemas, verificou-se que a solução obtida depende da forma como a região é percorrida, e portanto da solução inicial utilizada, o que indica a existência de diferentes condições de operação que correspondem ao mesmo custo total e ao mesmo caudal dos produtos  $G$  e  $H$ . Este facto foi confirmado utilizando o algoritmo OQNLP na resolução do problema em todos os pontos interiores. Na Figura 3.8 é apresentada a amplitude possível para os valores da variável  $u_9$  que correspondem ao mesmos valores do custo total e do caudal de  $G$ . Os resultados para as restantes variáveis são apresentados no Apêndice A.

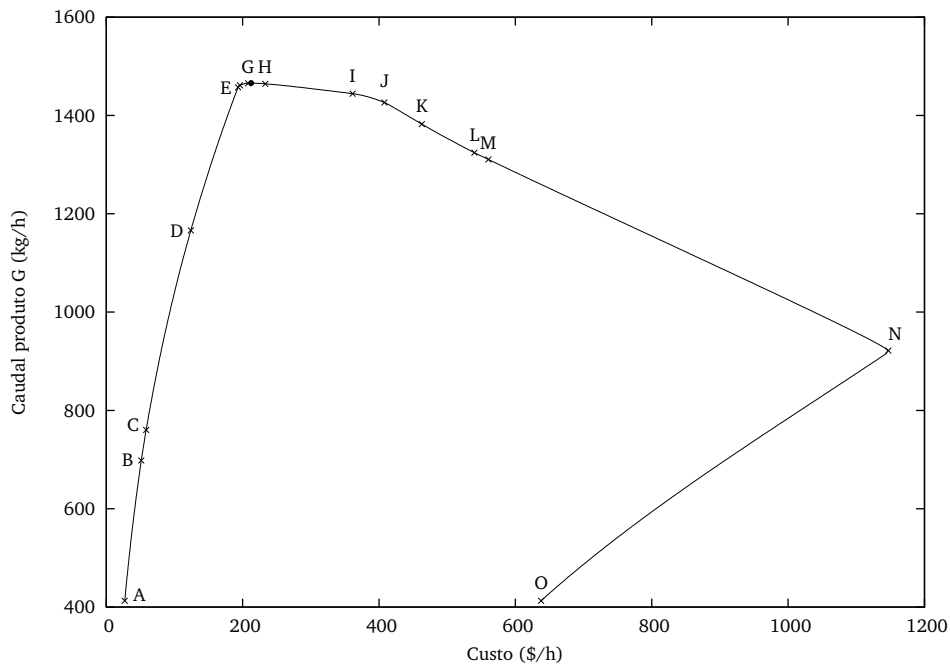


**Figura 3.8** Amplitude de variação da posição da válvula de vapor do *stripper* ( $u_9$ ) em função do custo total e do caudal mássico de  $G$ , no regime de operação  $50G/50H$ .

Pode verificar-se que, para esta variável, com a exceção dos pontos que constituem a fronteira, a amplitude de valores que correspondem ao mesmo estado estacionário é bastante grande. Mesmo na região próxima das condições correspondentes ao custo mínimo de operação, é possível obter o mesmo custo e o mesmo caudal de produto com valores desta variável que diferem quase 100%. Com maior ou menor amplitude, este comportamento é representativo do das restantes variáveis manipuláveis, em todos os regimes de operação. A exceção são as variáveis correspondentes ao caudais de alimentação dos reagentes  $D$ ,  $E$  e  $C$ , cuja amplitude de variação é bastante pequena, uma vez que são estes que determinam, em grande parte, a composição e o caudal do produto final.

De forma a melhor caracterizar as regiões de operação, nomeadamente as suas fronteiras, foram identificados os conjuntos de restrições activas. Para o regime de operação  $10G/90H$ , os resultados são apresentados na Figura 3.9. O conjunto de restrições activas em cada um dos segmentos são apresentados na Tabela 3.13. Os pontos representados na Figura foram obtidos por interpolação entre os pontos extremos das regiões apresentadas na Tabela. Para não sobrecarregar a figura, não são identificados todos os pontos. Todos os pontos apresentados foram obtidos utilizando o algoritmo OQNL, podendo existir diferentes conjuntos de restrições activas para um mesmo ponto.

Para este regime de operação, serão sobretudo de interesse as condições de operação nos segmentos  $DE$  e  $EF$ , com caudais próximos do valor máximo, mas com um custo



**Figura 3.9** Restrições activas nas condições de operação correspondentes aos custos máximo e mínimo no modo de operação 10G/90H.

total ligeiramente mais baixo. A variável  $u_2$  (caudal de alimentação de  $E$ ) encontra-se no seu limite superior em todas os segmentos próximos do caudal máximo, o que seria de esperar, uma vez que é esta variável que determina o caudal de  $H$  no produto final e que, neste regime, limita o caudal máximo possível. Na transição do custo mínimo para a produção máxima verifica-se que a variável  $u_9$  (caudal de vapor do *stripper*) passa do seu limite inferior para o limite superior, assumindo valores intermédios no segmento  $FG$ . Esta transição deve-se, por um lado à contribuição que esta variável tem para o custo de operação e, por outro, à necessidade de elevados caudal de vapor para conseguir a separação de grandes quantidades de produto, como se pode verificar na Tabela 3.11.

Os resultados obtidos para o regime de operação 50G/50H são apresentados na Figura 3.10 e na Tabela 3.14.

Para este regime, a região de maior interesse será o segmento  $CD$  que permite atingir caudais de produto próximos do máximo, com um custo significativamente mais baixo. Entre o ponto  $D$  e o ponto  $I$ , correspondente ao caudal máximo, a diferença entre o caudais de produto é, aproximadamente, 0.2%, enquanto que a diferença entre o custos é superior a 12%. Na transição entre o custo mínimo e a produção máxima, o caudal de alimentação de  $D$ ,  $u_1$ , atinge o seu limite superior, sendo este o reagente que limita o caudal de produto neste regime de operação. Novamente se verifica que o caudal de vapor do *stripper* passa do limite inferior para o limite superior, mantendo-se neste limite apenas numa região muito estreita em torno do caudal

**Tabela 3.13** Restrições activas nas condições de operação correspondentes aos custos mínimo e máximo no regime de operação 10G/90H.

Região	Custo	Caudal G	Restrições								
			$P_{react}^{max}$	$L_{react}^{min}$	$u_9^{min}$	$u_2^{max}$	$u_3^{max}$	$u_5^{max}$	$u_6^{max}$	$u_9^{max}$	$u_{11}^{max}$
<b>Custo mínimo</b>											
AB	27.05	412.77	1	1	1				1		
	50.95	694.83	1	1	1				1		
BC	51.68	701.71	1	1	1				1		1
	57.98	756.74	1	1	1				1		1
CD	58.83	763.62	1	1	1						1
	123.23	1162.63	1	1	1						1
DE	124.64	1169.51	1	1	1						
	193.08	1456.90	1	1	1						
EF	193.35	1457.90	1	1	1	1					
	195.08	1460.90	1	1	1	1					
FG	196.79	1461.90	1	1		1					
	203.68	1464.90	1	1		1					
<b>Custo máximo</b>											
GH	212.34 <sup>a</sup>	1465.90 <sup>a</sup>	1	1		1					1
	228.58	1464.90	1	1		1					1
HI	237.85	1463.90	1	1		1				1	1
	358.49	1444.90	1	1		1				1	1
IJ	363.98	1443.90	1	1		1			1	1	1
	406.95	1426.90	1	1		1			1	1	1
JK	408.66	1425.90	1	1		1	1		1	1	1
	462.13	1382.90	1	1		1	1		1	1	1
KL	463.34	1381.90	1	1		1			1	1	1
	534.97	1327.74	1	1		1			1	1	1
LM	544.90	1320.86	1	1		1			1		1
	555.22	1313.98	1	1		1			1		1
MN	565.59	1307.10	1	1	1	1			1		1
	1147.08	921.85	1	1	1	1			1		1
NO	1143.30	914.97	1	1	1				1		1
	637.56	412.77	1	1	1				1		1

<sup>a</sup> produção máxima

máximo. Também o caudal da água de arrefecimento do separador,  $u_{11}$ , atinge o seu limite superior. A válvula de reciclo do compressor,  $u_5$ , abandona o limite inferior.

Os resultados para o regime de operação 90G/10H são apresentados na Figura 3.11 e na Tabela 3.15.

As características deste regime de operação são semelhantes às do regime 50G/50H, sendo que aqui a região de maior interesse será a representada pelo segmento CD. A operação nesta região permite caudais praticamente iguais ao máximo, com uma diferença inferior a 0.2%, com um custo cerca de 19% inferior. As principais diferenças estão no comportamento da variável  $u_{11}$  (caudal da água de arrefecimento do condensador), que atinge o seu limite superior ainda na zona do custo mínimo. Novamente



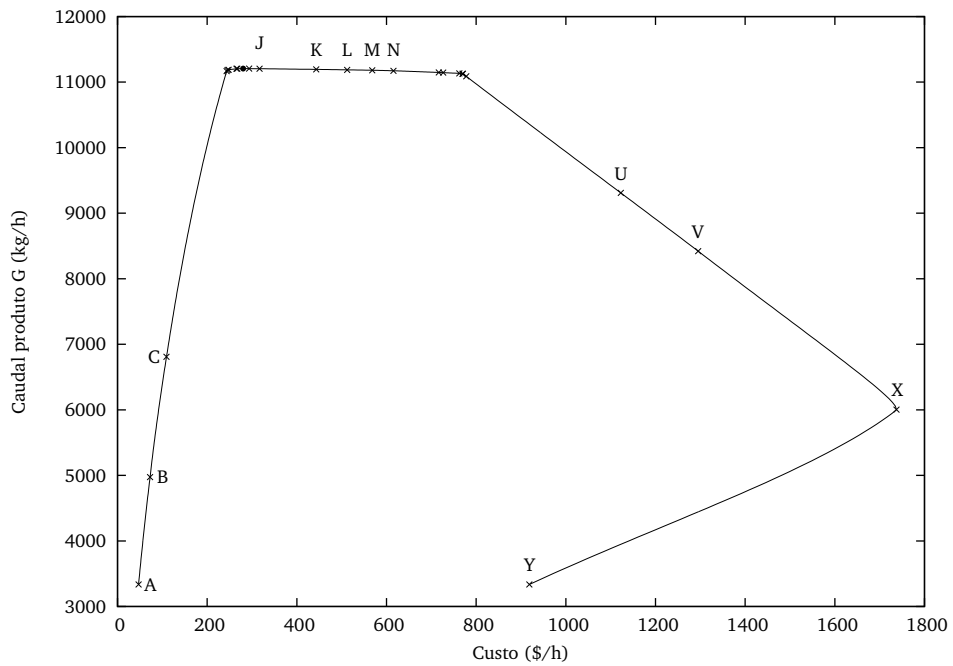


Figura 3.10 Restrições activas nas condições de operação correspondentes aos custos máximo e mínimo no modo de operação 50G/50H.

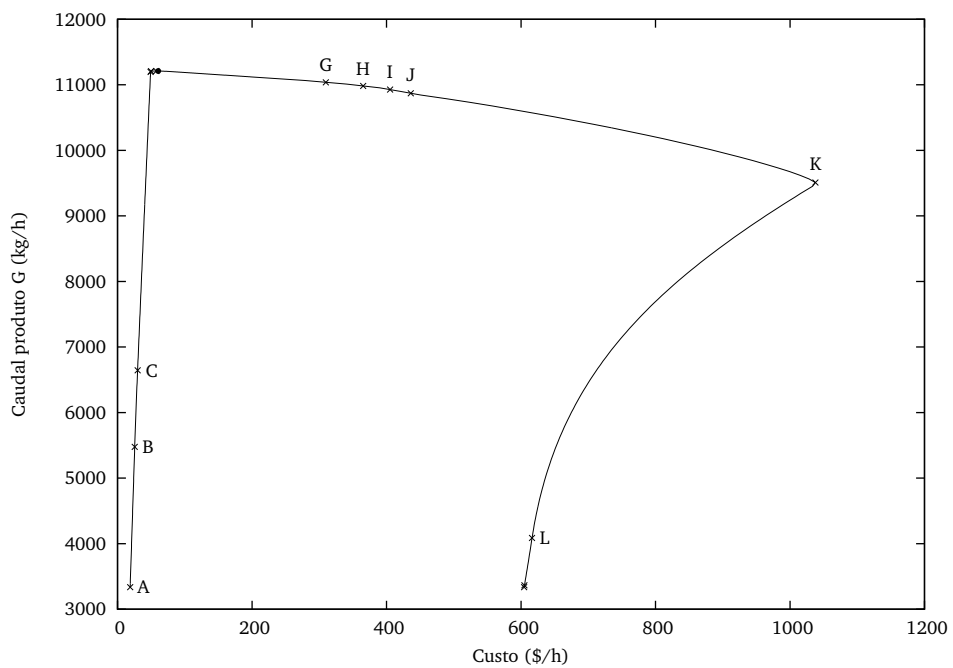


Figura 3.11 Restrições activas nas condições de operação correspondentes aos custos máximo e mínimo no modo de operação 90G/10H.

**Tabela 3.14** Restrições activas nas condições de operação correspondentes aos custos mínimo e máximo no regime de operação 50G/50H.

Região	Custo	Caudal G	Restrições								
			$p_{react}^{max}$	$L_{react}^{min}$	$u_5^{min}$	$u_9^{min}$	$u_1^{max}$	$u_2^{max}$	$u_3^{max}$	$u_6^{max}$	$u_9^{max}$
<b>Custo mínimo</b>											
AB	46.94	3334.88	1	1	1	1					
	72.16	4946.73	1	1	1	1					
BC	73.11	5002.31	1	1		1					
	108.62	6780.91	1	1		1					
CD	109.90	6836.50	1	1	1	1					
	243.44	11168.25	1	1	1	1					
DE	243.48	11169.25	1	1	1	1	1				
	245.21	11184.25	1	1	1	1	1				
EF	245.87	11185.25	1	1	1		1				
	247.47	11187.25	1	1	1		1				
FG	248.39	11188.25	1	1			1				
	264.64	11203.25	1	1			1				
GH	266.05	11204.25	1	1			1			1	
HI	268.34	11205.25	1	1			1			1	1
	280.07 <sup>a</sup>	11206.25 <sup>a</sup>	1	1			1			1	1
<b>Custo máximo</b>											
IJ	306.86	11205.25	1	1			1				1
JK	326.69	11204.25	1	1		1	1				1
	438.12	11194.25	1	1		1	1				1
KL	448.35	11193.25	1	1	1	1	1				1
	507.56	11187.25	1	1	1	1	1				1
LM	517.11	11186.25	1	1		1	1				1
	563.69	11181.25	1	1		1	1				1
MN	572.72	11180.25	1			1	1	1			1
	612.76	11173.25	1			1	1	1			1
NO	618.20	11172.25	1	1		1	1	1			1
	714.75	11148.25	1	1		1	1	1			1
OP	718.26	11147.25	1			1	1	1			1
	724.81	11145.25	1			1	1	1			1
PQ	728.06	11144.25	1				1	1			1
	761.28	11133.25	1				1	1			1
QR	763.31	11132.25	1				1	1	1		1
	768.86	11128.25	1				1	1	1		1
RS	770.09	11127.25	1	1			1	1	1		1
ST	770.62	11126.25	1	1			1	1	1	1	1
	772.53	11116.25	1	1			1	1	1	1	1
TU	783.15	11060.67	1	1				1	1	1	1
	1117.11	9337.65	1	1				1	1	1	1
UV	1127.96	9282.07	1	1		1		1	1	1	1
	1289.78	8448.35	1	1		1		1	1	1	1
VX	1300.41	8392.77	1	1		1		1		1	1
	1737.50	6002.78	1	1		1		1		1	1
XY	1727.81	5947.20	1	1		1			1		1
	918.59	3334.88	1	1		1			1		1

<sup>a</sup> produção máxima

**Tabela 3.15** Restrições activas nas condições de operação correspondentes aos custos mínimo e máximo no regime de operação 90G/10H.

Região	Custo	Caudal G	Restrições							
			$P_{react}^{max}$	$L_{react}^{min}$	$u_9^{min}$	$u_1^{max}$	$u_5^{max}$	$u_6^{max}$	$u_9^{max}$	$u_{11}^{max}$
<b>Custo mínimo</b>										
AB	18.76	3336.43	1	1	1		1			
	25.47	5449.50	1	1	1		1			
BC	25.64	5505.10	1	1	1		1			1
	29.74	6617.25	1	1	1		1			1
CD	29.97	6672.85	1	1	1					1
	49.18	11196.42	1	1	1					1
DE	49.18	11197.42	1	1	1	1				1
	49.23	11208.42	1	1	1	1				1
EF	50.36	11209.42	1	1		1				1
	54.93	11210.42	1	1		1				1
<b>Custo máximo</b>										
FG	60.49 <sup>a</sup>	11211.42 <sup>a</sup>	1	1		1			1	1
	278.88	11065.81	1	1		1			1	1
GH	340.78	11010.21	1			1			1	1
HI	389.79	10954.60	1	1		1	1		1	
IJ	420.97	10898.99	1	1		1	1	1	1	
JK	451.04	10843.38	1	1	1	1		1		1
	1037.71	9508.81	1	1	1	1		1		1
KL	1033.52	9453.21	1	1	1			1		1
	616.88	4114.93	1	1	1			1		1
LM	616.04	4059.32	1	1	1		1	1		1
	605.56	3392.03	1	1	1		1	1		1
MN	604.63	3336.43	1	1	1		1	1		

<sup>a</sup> produção máxima

se verifica que, na transição do custo mínimo para a produção máxima, o caudal de vapor do *stripper*,  $u_9$ , passa do seu valor mínimo para o valor máximo.

É importante notar que, para todos os regimes, a operação nas condições de custo mínimo ou de produção máxima implica que a pressão do reactor esteja no seu limite superior e o nível no seu limite inferior. Também se verifica que existem variáveis com efeitos contrários sobre o custo de operação e o caudal de produto. O melhor exemplo de tal é o caudal de vapor do *stripper* que, na região do custo mínimo, assume sempre o seu valor mínimo, enquanto que para obter o caudal máximo de produto é necessário que esteja no seu limite superior.

Alguns dos resultados coincidem com os apresentados por Ricker (1995) para o caudal máximo de produto, nomeadamente o ponto *F*, no regime 10G/90H, e os pontos *E*, nos regimes 50G/50H e 90G/10H. Estes pontos são o limite superior dos segmentos da fronteira correspondente ao custo mínimo em que a alimentação do reagente limitante atinge o seu limite superior, mas em que o caudal de vapor do *stripper* ainda

se encontra no seu limite inferior.

### 3.4.4 Ganhos estacionários

De forma semelhante à análise feita ao comportamento do reactor, na Secção 3.3.3, foram obtidos os ganhos estacionários das variáveis medidas relativamente às variáveis manipuláveis do processo. Novamente, os resultados foram obtidos determinando o valor, em estado estacionário, das variáveis do processo, para diferentes valores das variáveis manipuláveis. Para tal, é resolvido o sistema de equações algébricas constituído pelas equações do modelo, considerando nulos os termos diferenciais e fixando as variáveis manipuláveis. Foi utilizando o modelo do processo em GAMS atrás descrito e, para a resolução do sistema de equações, a formulação CNS (*constrained nonlinear systems*) (Rosenthal, 2010).

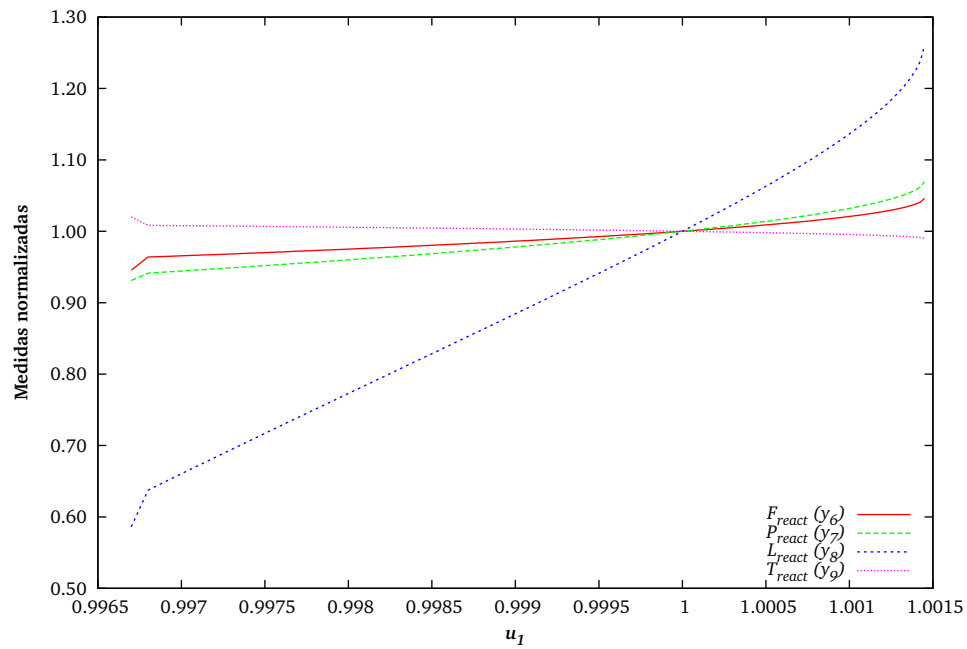
A utilização desta formulação implica que o sistema de equações algébricas tenha o mesmo número de variáveis e de equações. Para que tal ocorra, foi necessário considerar como variáveis o caudal de líquido do separador ( $u_7$ ) e o caudal de líquido do *stripper* ( $u_8$ ), que são determinadas pela solução do sistema de equações, ao contrário das restantes variáveis manipuláveis, cujo valor é estabelecido à partida. Estas variáveis foram seleccionadas por se ter verificado que, com valores arbitrários destas, não é possível satisfazer o balanço mássico, em estado estacionário, a estas unidades. Os limites considerados para as variáveis são agora os os limites de segurança apresentados na Tabela 3.6.

Os resultados obtidos foram também utilizados para verificar a coincidência do modelo utilizado com o modelo original, introduzindo alguns dos resultados obtidos neste último. Verificou-se que os estados estacionários obtidos com o modelo modificado correspondem também a estados estacionários do modelo original.

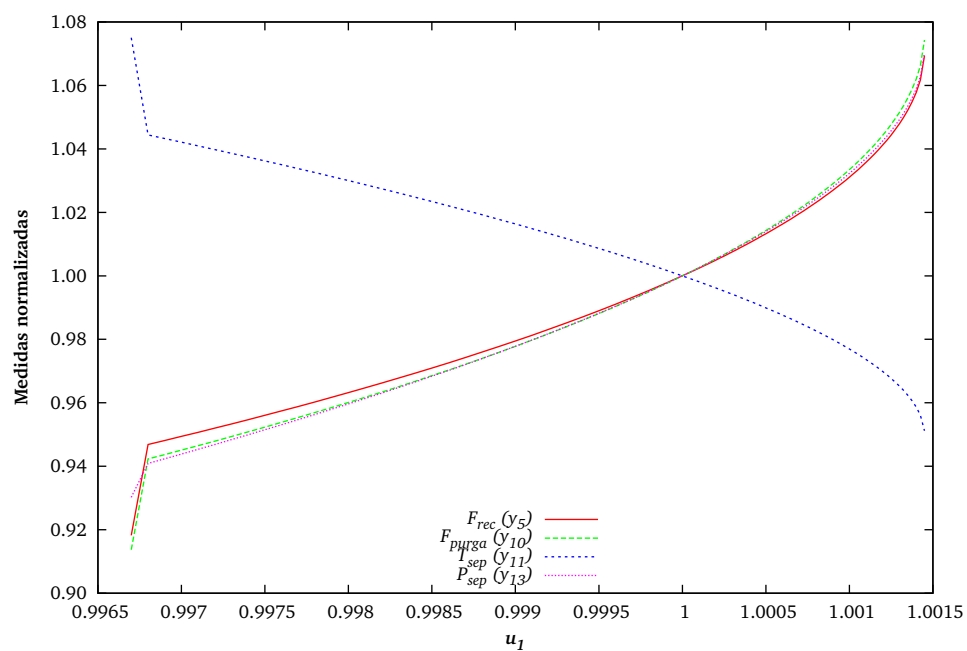
Para esta discussão foi considerada a variação das variáveis manipuláveis em torno dos valores apresentados na Tabela 3.4, correspondente ao modo de operação 1 da Tabela 3.1. Apenas são apresentados os resultados que se consideram mais relevantes para a caracterização do processo. A totalidade dos resultados obtidos é apresentada no Apêndice B.

Na Figura 3.12 são apresentados os resultados obtidos para o caudal de alimentação de  $D$ ,  $u_1$ . Nestes resultados, e nos seguintes, todas as variáveis são normalizadas utilizando os valores correspondentes ao modo base de operação (modo 1), apresentados nas Tabelas 3.3, 3.4 e 3.5.

O intervalo de variação bastante pequeno deve-se ao facto de este caudal ter uma magnitude bastante elevada no estado estacionário de referência (3664 kg/h). O mesmo se verifica para os resultados obtidos para os restantes caudais de alimentação, com

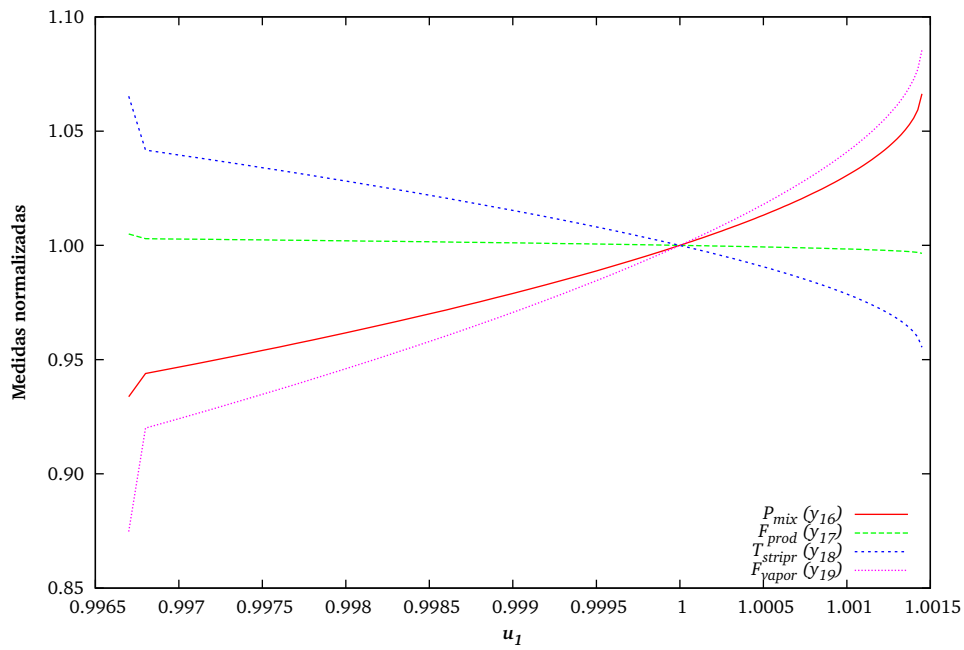


(a)

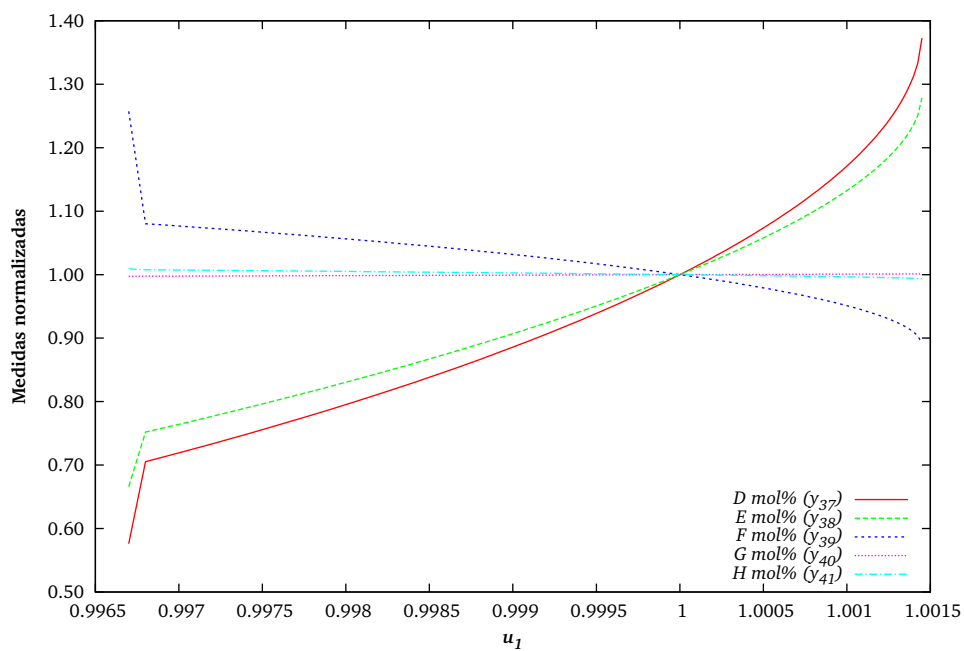


(b)

**Figura 3.12** Variáveis medidas em função do caudal de alimentação de  $D$  ( $u_1$ ).



(c)



(d)

**Figura 3.12 (cont.)** Variáveis medidas em função do caudal de alimentação de  $D$  ( $u_1$ ).

excepção do caudal de alimentação de  $A$ , cujo valor no estado estacionário de referência é significativamente menor (22.4 kg/h). É importante referir que as posições das válvulas que determinam estes caudais são definidas em percentagem, pelo que uma pequena alteração na posição da válvula provoca uma variação significativa do valor absoluto do caudal.

Em todos os resultados apresentados na Figura 3.12 se verifica que o caudal da alimentação de  $D$  tem uma grande influência sobre o estado estacionário do processo. No caso do nível do reactor, uma variação negativa de 0,35% provoca uma diminuição de 40% no nível do reactor. Também se pode verificar que, na região em torno do modo base de operação, os ganhos se mantêm aproximadamente constantes. Já nos extremos do intervalo de variação ocorrem mudanças significativas do ganho de todas as variáveis apresentadas. Esta variação é particularmente notória no limite inferior do intervalo de variação considerado. A inflexão observada em todas as curvas deve-se à alteração da superfície de transferência de calor no reactor que, nesse ponto, passa a ser inferior à área total disponível. Nas Figuras 3.12a, 3.12b e 3.12c podemos verificar que, no limite superior do intervalo, os ganhos do nível do reactor ( $y_8$ ), da pressão do separador ( $y_{13}$ ) e da pressão do tanque de mistura ( $y_{16}$ ) variam muito rapidamente, tendendo, em valor absoluto, para infinito.

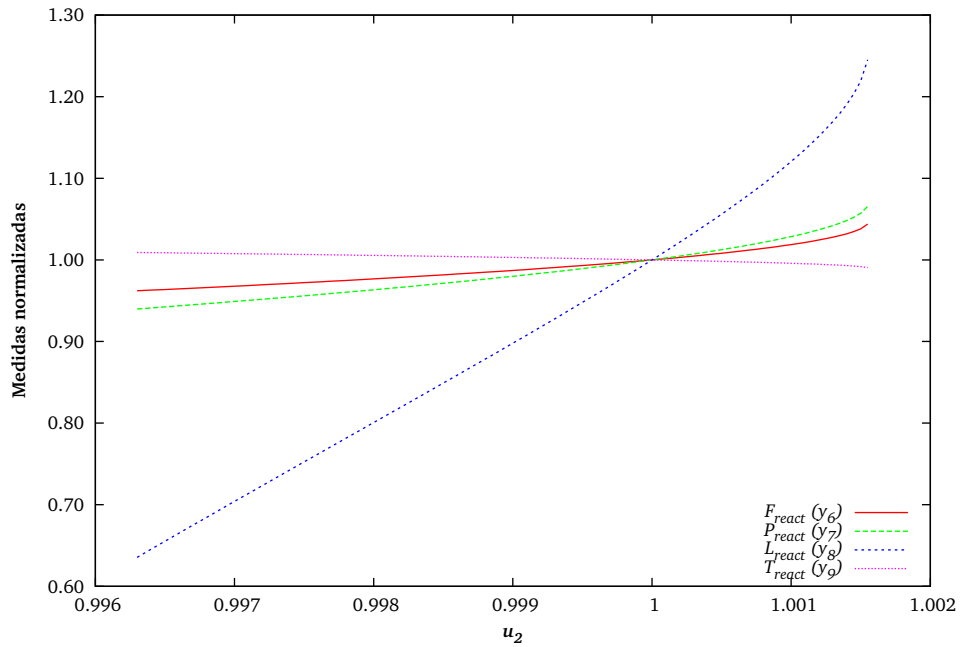
É ainda de referir que, apesar de as quantidades totais da fase líquida no separador e no *stripper* não serem variáveis, se observam variações do nível nessas unidades, que se devem à variação da temperatura e da composição e, em consequência, da densidade da fase líquida.

Na Figura 3.13 são apresentados os resultados para o caudal de alimentação de  $E$  ( $u_2$ ).

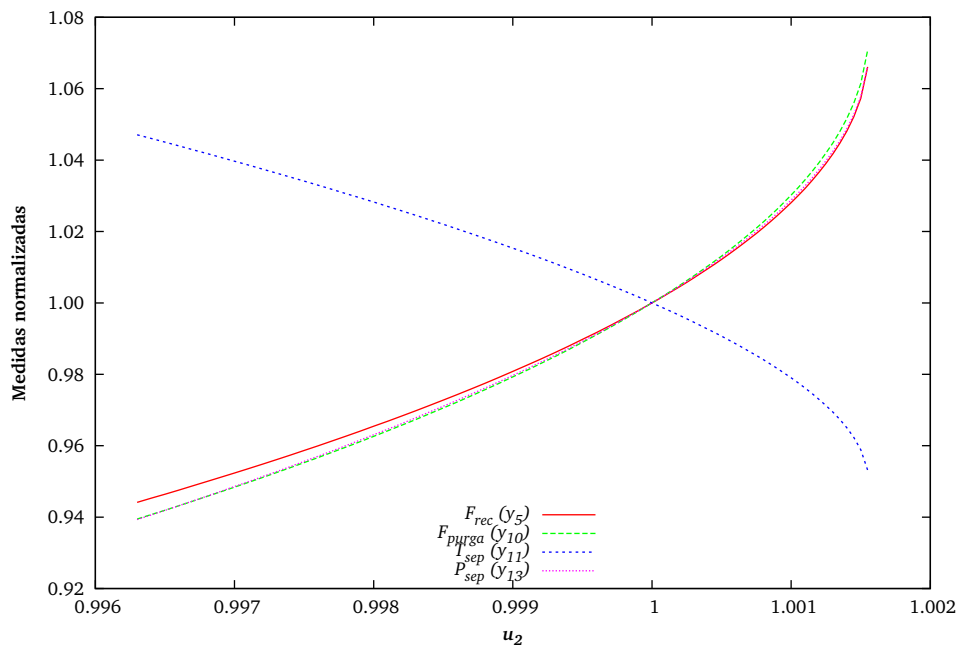
Mais uma vez, se observa que, para pequenas variações da posição da válvula, se verificam grandes variações nas variáveis medidas. No caso do nível do reactor, esta variação é de -35% a 25%, aproximadamente, para uma variação da posição da válvula entre -0,4% e 0,2%. De notar que aqui não se observa a súbita alteração dos ganhos, observada com o caudal de alimentação de  $D$  quando a superfície de transferência no reactor começa a variar. Tal deve-se ao facto de não ter sido possível obter estados estacionários na zona de operação correspondente a essa condição. Mais uma vez se verifica que, no limite superior do intervalo, os ganhos da pressão do tanque de mistura, da pressão do separador e do nível do reactor aumentam muito rapidamente, tendendo para infinito.

Na Figura 3.14 são apresentados os resultados obtidos para o caudal de vapor no *stripper*,  $u_9$ .

Observa-se uma variação bastante significativa dos ganhos estacionários da generalidade das variáveis medidas. Estas variações são particularmente importantes para valores próximos de 0,2, onde ocorre a inflexão devida à alteração da área de trans-



(a)



(b)

Figura 3.13 Variáveis medidas em função do caudal de alimentação de E ( $u_2$ ).



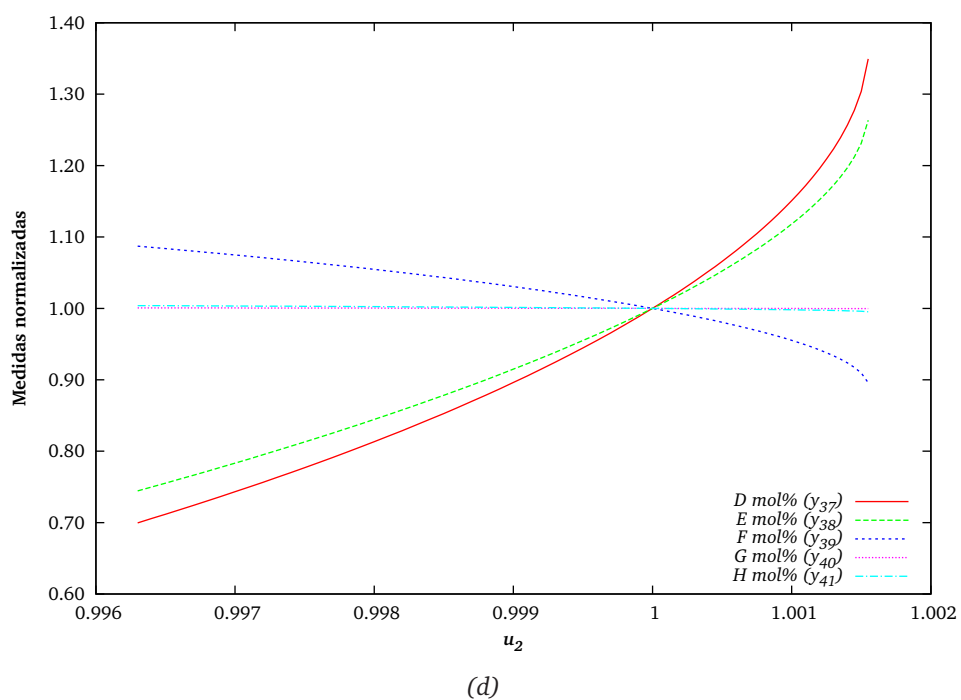
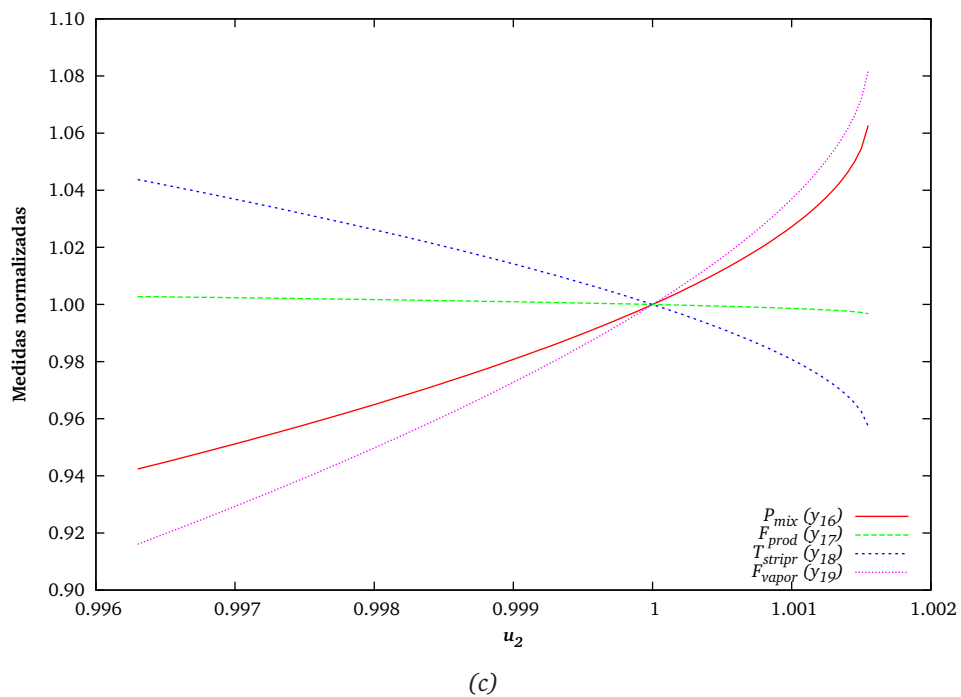
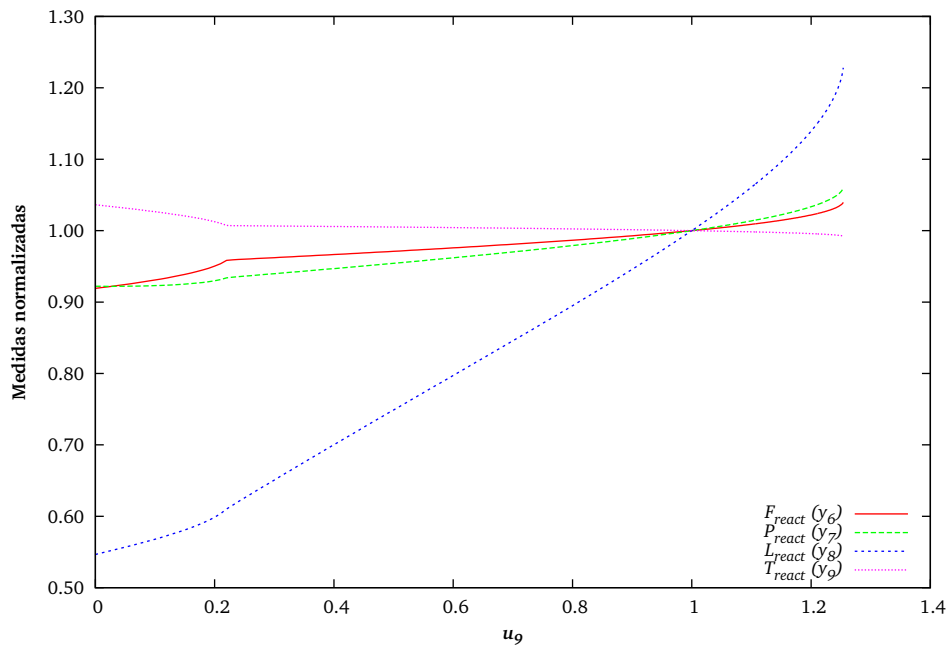
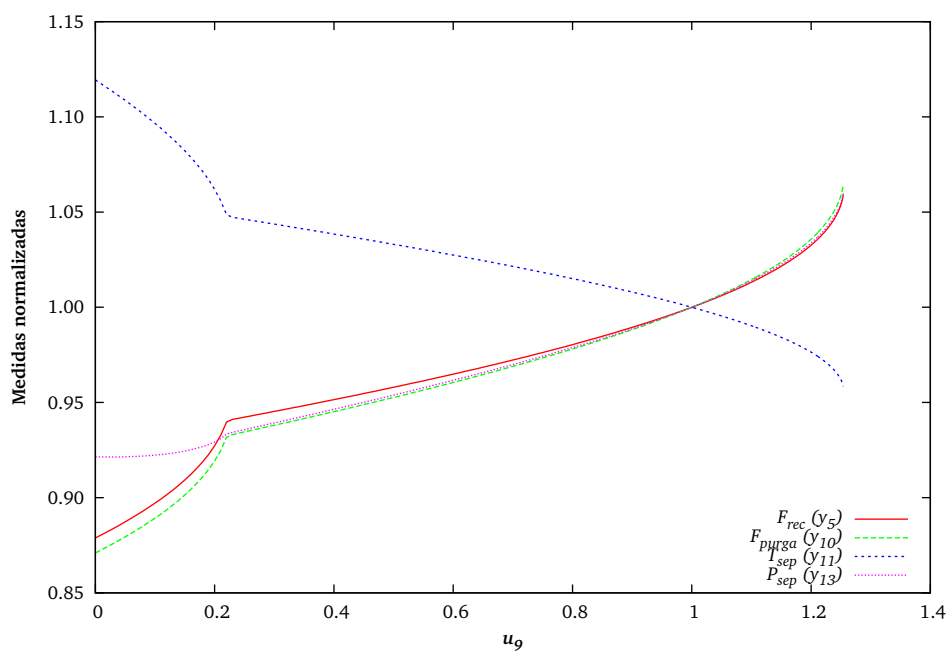


Figura 3.13 (cont.) Variáveis medidas em função do caudal de alimentação de  $E$  ( $u_2$ ).



(a)



(b)

Figura 3.14 Variáveis medidas em função do caudal de vapor no *stripper* ( $u_9$ ).

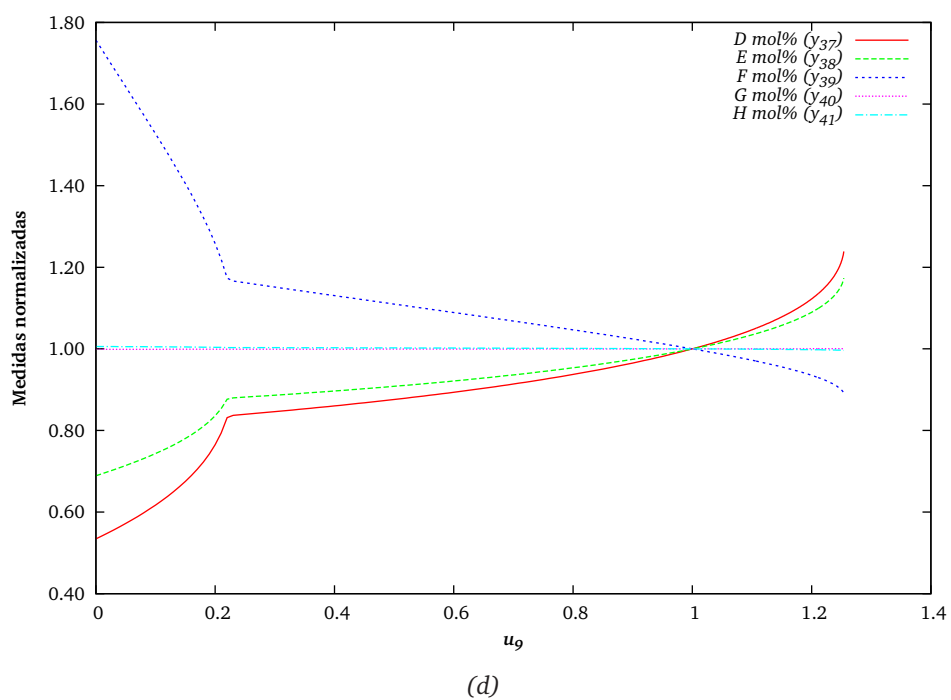
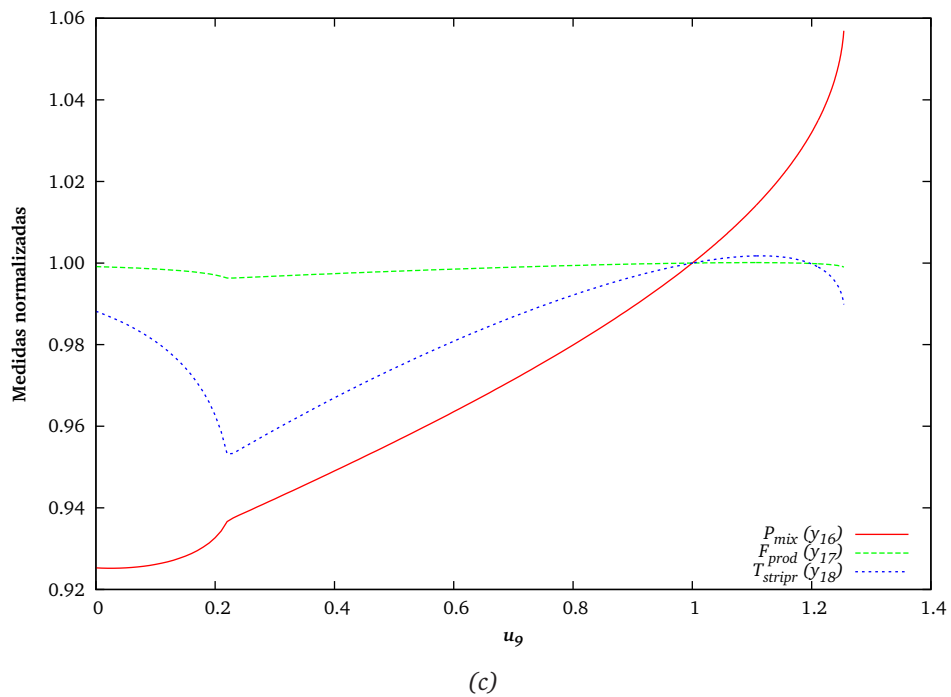


Figura 3.14 (cont.) Variáveis medidas em função do caudal de vapor no *stripper* ( $u_9$ ).

ferência de calor no reactor, e no limite superior do intervalo, quando alguns ganhos aumentam muito rapidamente. É particularmente relevante o comportamento da temperatura do *stripper* ( $y_{17}$ ), apresentado na Figura 3.14c. O ganho desta variável muda de sinal, sendo negativo até valores próximos de 0,2, e positivo a partir daí, até valores próximos de 1,2, quando volta a ser negativo. Nesta última transição existe ainda uma região em que o ganho é praticamente nulo. Como consequência, o mesmo valor da temperatura do *stripper* pode ser conseguido com dois valores distintos do caudal de vapor.

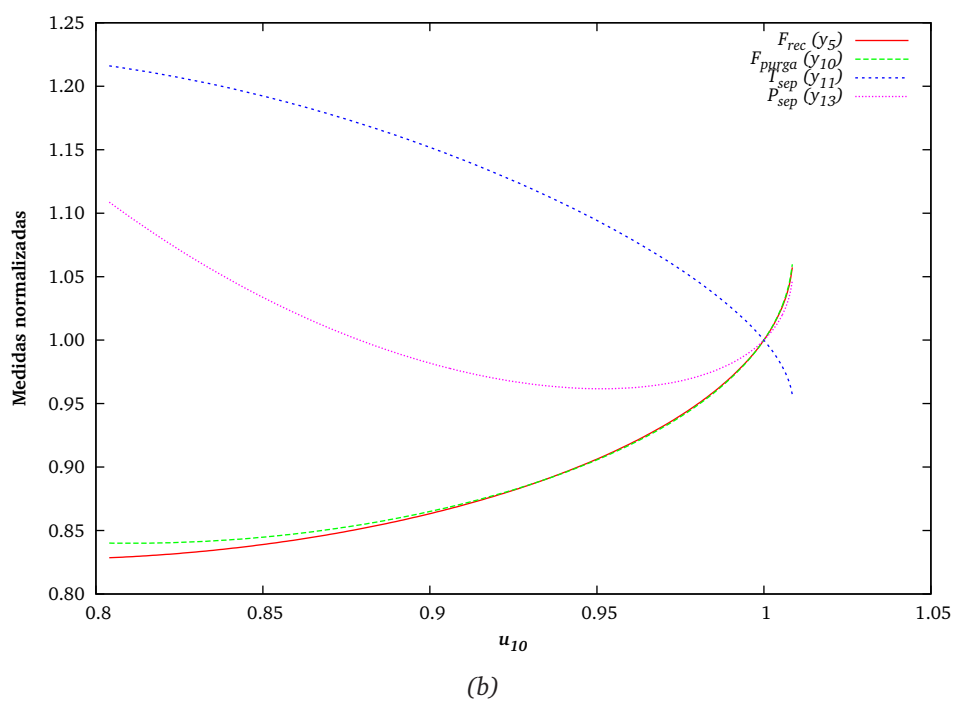
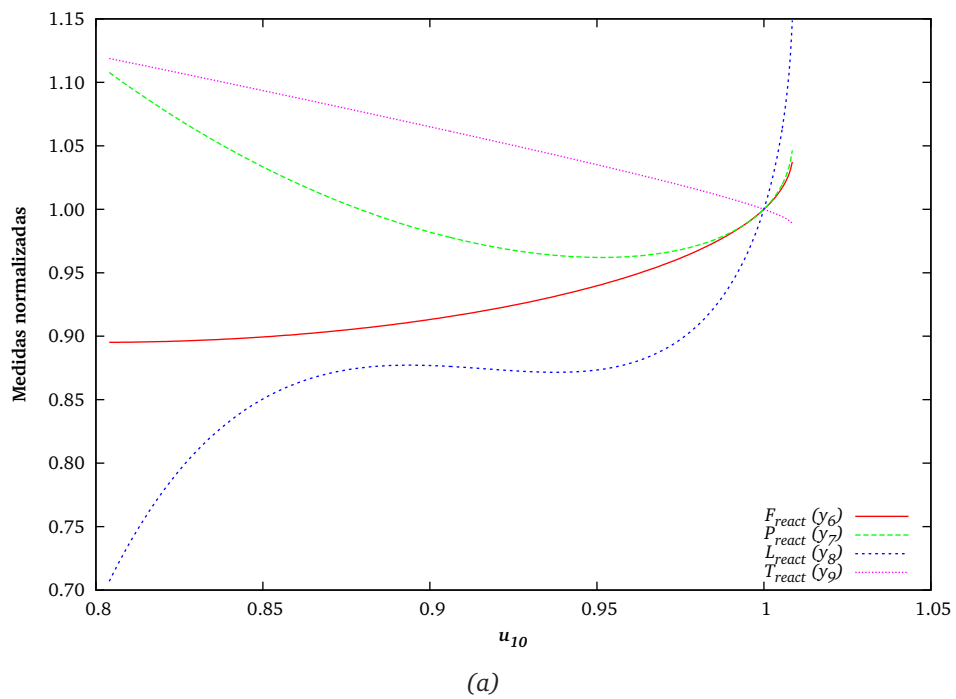
Na Figura 3.15 são apresentados os resultados relativos ao caudal de água de arrefecimento do reactor,  $u_{10}$ .

Podemos observar que o ganho estacionário de um grande conjunto de variáveis medidas muda muito rapidamente no extremo superior do intervalo, tendendo para infinito. Podemos ainda observar mudanças no sinal do ganho de algumas variáveis, como a pressão e o nível do reactor ( $y_7$  e  $y_8$  na Figura 3.15a), a pressão do separador ( $y_{13}$  na Figura 3.15b) e a pressão do tanque de mistura ( $y_{16}$  na Figura 3.15c). Estas mudanças de sinal fazem com que, novamente, ocorra multiplicidade de entradas. Esta situação é particularmente notória no comportamento do nível do reactor, para o qual existem valores que correspondem a quatro valores distintos do caudal da água de arrefecimento. Observa-se ainda que o ganho do nível é inicialmente muito elevado, passando depois para uma região em que a curva é quase plana (e, por isso, o seu ganho praticamente nulo), e, finalmente, aumentando muito rapidamente no extremo superior do intervalo. De notar também que o aumento muito rápido do ganho estacionário do nível do reactor, bem como das restantes variáveis, ocorre para valores do caudal da água de arrefecimento muito próximos dos correspondentes ao estado estacionário de referência.

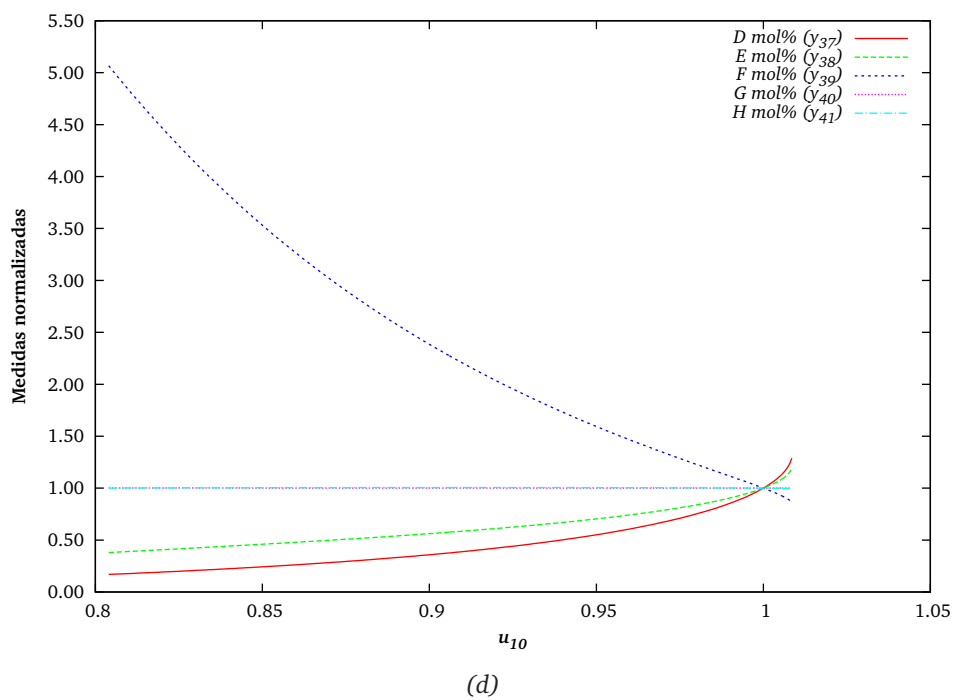
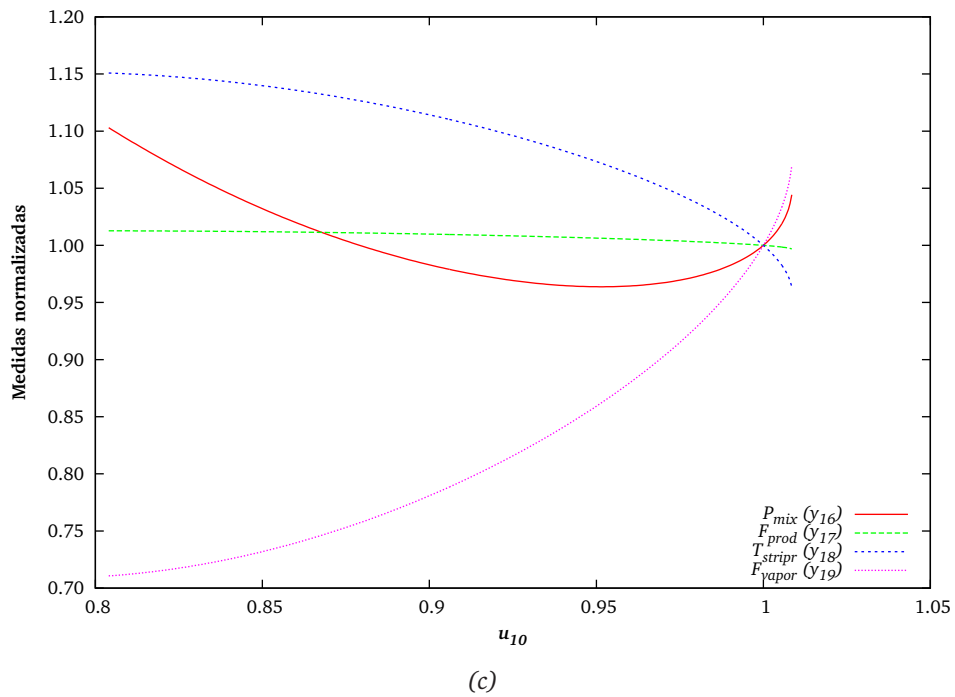
Na Figura 3.16 são apresentados os resultados obtidos para o caudal da água de arrefecimento do separador,  $u_{11}$ .

Também para esta variável manipulada se observa a mudança do sinal do ganho e a existência de multiplicidade de entradas, nomeadamente das pressões do separador ( $y_{12}$  na Figura 3.16b) e do tanque de mistura ( $y_{16}$  na Figura 3.16c). Observa-se também uma variação muito rápida da generalidade dos ganhos para valores do caudal da água de arrefecimento próximos dos correspondentes ao estado estacionário de referência.

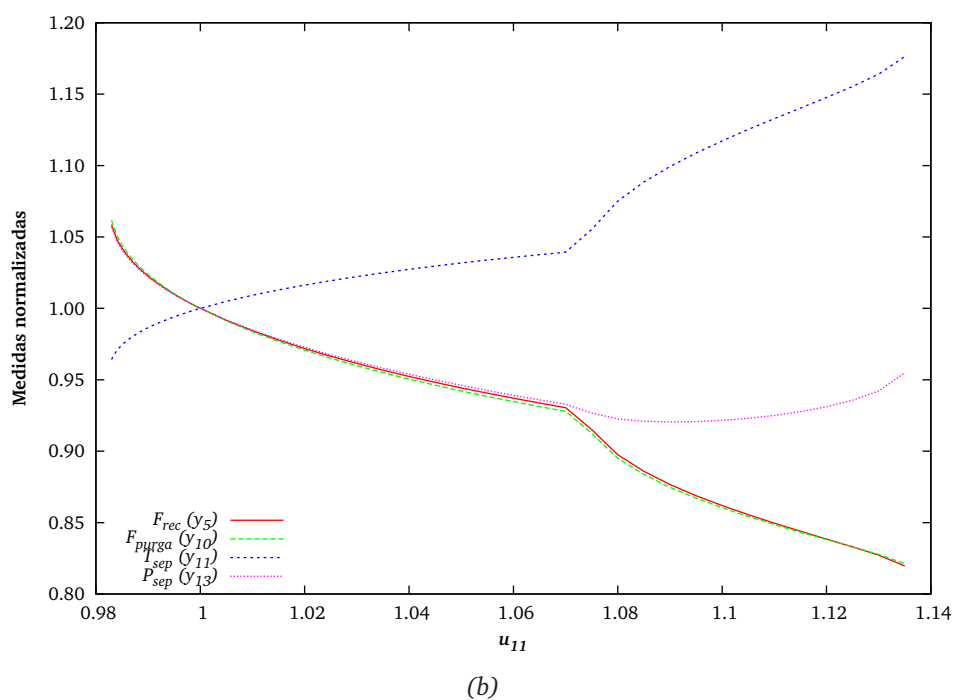
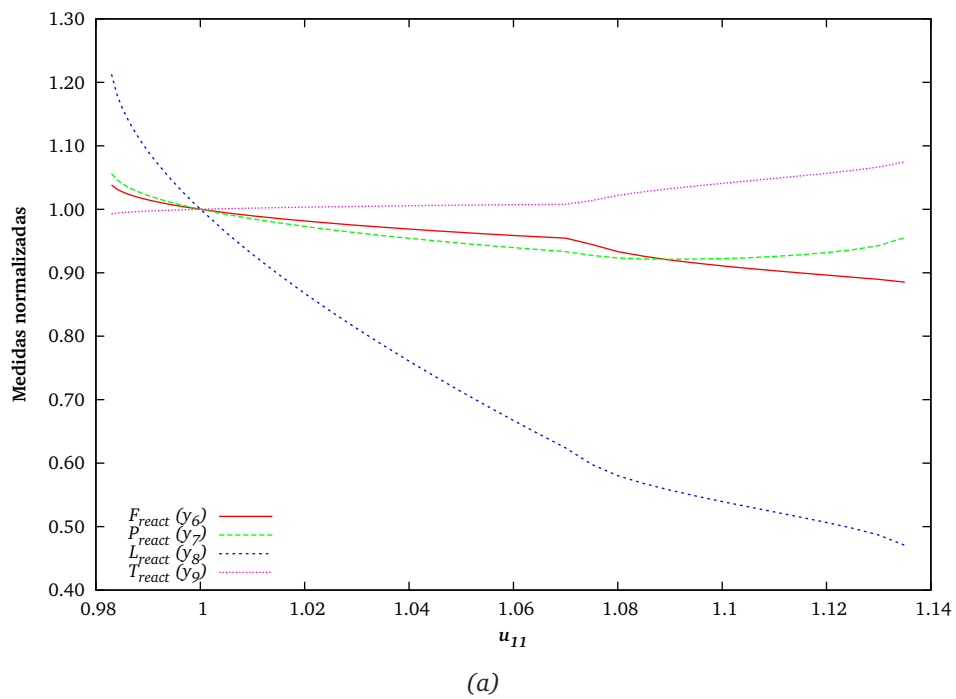
Dos resultados apresentados, por comparação com os resultados obtidos para o reactor quando considerado isoladamente, confirma-se a forte interacção existente entre as unidades, sendo o comportamento do reactor, quando integrado no processo, totalmente distinto de quando considerado isoladamente. Estas diferenças são particularmente visíveis nos gráficos dos ganhos relativos aos caudais das alimentações (Figuras 3.5, para o reactor, e 3.12 e 3.13, para o processo). Enquanto que para o re-



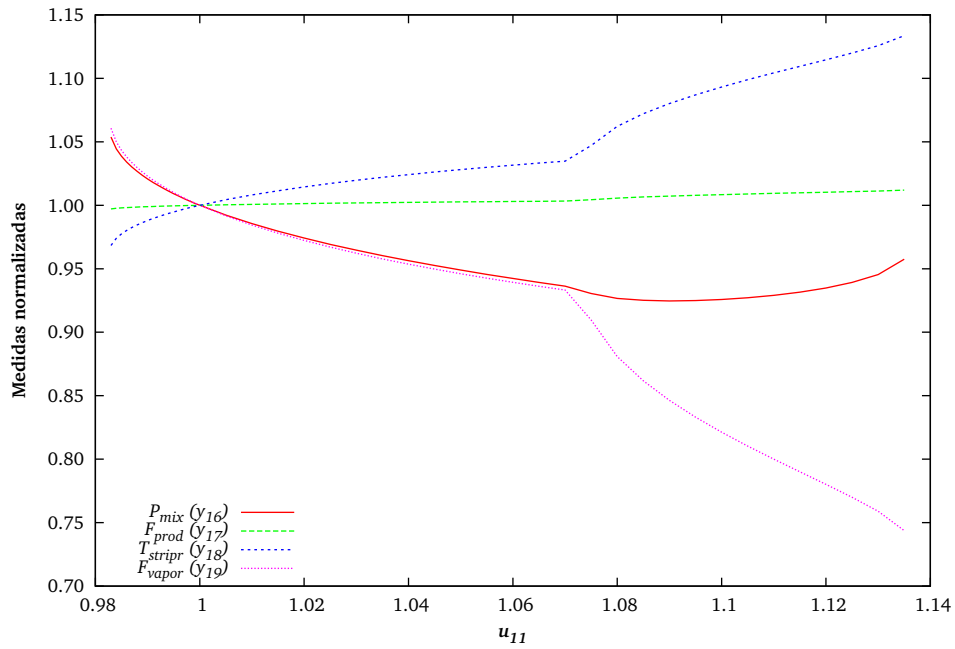
**Figura 3.15** Variáveis medidas em função do caudal de água de arrefecimento do reactor ( $u_{10}$ ).



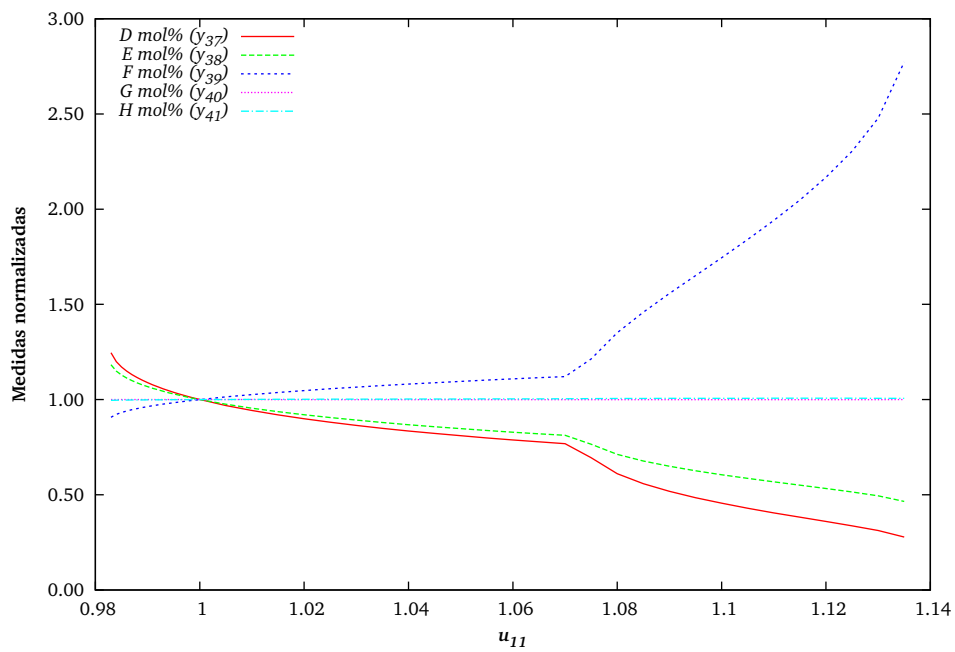
**Figura 3.15 (cont.)** Variáveis medidas em função do caudal de água de arrefecimento do reactor ( $u_{10}$ ).



**Figura 3.16** Variáveis medidas em função do caudal de água de arrefecimento do separador ( $u_{11}$ ).



(c)



(d)

Figura 3.16 (cont.) Variáveis medidas em função do caudal de água de arrefecimento do separador ( $u_{11}$ ).



actor isoladamente o valor dos caudais das alimentações tem pouca influência sobre as condições de operação do reactor, essa influência é muito grande quando o reactor se encontra integrado no processo. A principal diferença entre as duas situações é a existência da corrente de reciclo. Quando o processo é considerado globalmente, uma pequena variação nos caudais das alimentações provoca grandes variações no caudal do reciclo e, através deste, nas condições de operação do reactor. Este efeito não ocorre quando o reactor é considerado isoladamente. Esta é mais uma razão pela qual não é adequado, para este processo, a utilização de uma estrutura de controlo baseada no controlo individual de cada unidade. Deverá sempre ser tida em conta a interacção entre as diferentes unidades e a natureza multivariável do processo.

É ainda importante referir que, em todos os resultados apresentados, os ganhos estacionários das fracções de  $G$  e  $H$  no produto final ( $y_{40}$  e  $y_{41}$ , respectivamente) são praticamente nulos relativamente a todas as variáveis medidas. Isto apesar de se verificarem grandes alterações relativas na composição do produto final em termos relativos, estas, em termos absolutos, são muito pequenas, uma vez que as fracções molares dos restantes compostos no estado estacionário de referências são muito pequenas. Tal deve-se ao facto de, em estado estacionário, quer a composição quer o caudal de produto serem sobretudo determinados pelos caudais das alimentações, mantendo-se por isso aproximadamente constantes quando os caudais das alimentações são fixos, como acontece nos resultados obtidos para as restantes variáveis manipuláveis.

### 3.5 Conclusões

Os resultados apresentados permitem retirar conclusões importantes sobre comportamento do processo. Verifica-se a existência de forte interacção entre as diferentes unidades, demonstrada pela diferença no comportamento do reactor quando considerado isoladamente e quando integrado no processo. Estas diferenças devem-se à corrente de reciclo que, na generalidade das condições de operação, constitui uma parte importante da mistura que alimenta o reactor (superior a 50% no modo base de operação). A interacção entre as diferentes unidades também pode ser observada nos resultados relativos aos ganhos estacionários, em que, por exemplo, uma descontinuidade na variação da área de transferência de calor no reactor provoca alterações significativas nos ganhos de todas as variáveis medidas.

Outras características importantes são a não-linearidade, demonstrada pelos resultados obtidos para os ganhos estacionários, a instabilidade em ciclo aberto e a existência de modos integrantes, demonstradas pela linearização do modelo. Estas características tornam complexa a aplicação de algoritmos de optimização, quer estacionária quer dinâmica. A existência de variáveis integrantes que levou à necessidade de

reformulação do modelo para resolver os problemas de maximização da produção, minimização do custo e determinação de estados estacionários.

A análise aos ganhos estacionários das variáveis medidas do processo relativamente às variáveis manipuláveis permite verificar que alguns destes ganhos variam significativamente numa região próxima das condições de operação correspondentes ao modo base de operação, ocorrendo mudanças do sinal de alguns destes ganhos. Também se verifica a grande sensibilidade do processo relativamente aos caudais das alimentações dos reagentes  $C$ ,  $D$  e  $E$ , observando-se grandes variações nas variáveis medidas para muito pequenas alterações nestas variáveis.

A realização da análise de ganhos estacionários permitiu também verificar que, para pequenas variações no caudais das alimentações de  $C$ ,  $D$  e  $E$ , não é possível a obtenção de novos estados estacionários mantendo inalteradas as restantes variáveis manipuláveis. Por outro lado, os resultados relativos às condições óptimas de operação mostram que estes caudais são determinantes para o caudal e a composição do produto. Isto implica que todas as variáveis manipuladas tenham que ser coordenadas com estas, tal como na estrutura proposta por Ricker (1996), que utiliza um esquema de *ratios* dependentes do caudal do produto para estabelecer referências para a generalidade dos controladores.

As variáveis mais críticas do processo são a pressão e o nível do reactor. A operação nas condições de custo mínimo, para qualquer caudal de produto, ou de produção máxima implicam que ambas as variáveis se encontrem nos seus limites, superior no caso da pressão, e inferior no caso do nível. Esta combinação de variáveis permite que a quantidade de material na fase gasosa do reactor seja o mais elevada possível, aumentando assim as taxas de reacção. Daqui se pode concluir que a operação do processo TE se encontra condicionada pelas dimensões do reactor. Se a sua dimensão fosse superior seriam possíveis maiores caudais de produto, sem que fosse necessário manter estas variáveis nos seus limites, sendo, por isso, mais simples o controlo do processo.

Os caudais das alimentações de  $C$ ,  $D$  e  $E$  dependem fortemente do caudal e da composição pretendidos para o produto final, de forma quase independente do custo de operação. O mesmo não se passa com as restantes variáveis manipuláveis. Ou seja, o mesmo caudal de produto, com a mesma composição, pode ser conseguido com diferentes combinações de valores das variáveis manipuláveis, resultando no mesmo custo de operação. Tal apenas não ocorre para as condições de operação correspondentes ao custo mínimo e à produção máxima, ainda que para regiões de operação próximas do custo mínimo, o mesmo caudal de produto pode ser conseguido para diferentes combinações de valores das variáveis manipuláveis, com o mesmo custo. Isso mesmo pode ser verificado na Figura 3.8, para o caudal de vapor do *stripper*, e nos restantes resultados apresentados no Anexo A. Nas condições de custo mínimo, ou produção máxima, apenas o caudal de água de arrefecimento e a velocidade de

agitação do reactor podem assumir diferentes valores (uma vez que diferentes combinações destas variáveis resultam na mesma quantidade de calor transferida).

A região do custo mínimo de operação e a região de produção máxima têm características diferentes, em qualquer um dos modos de operação. Diferem, nomeadamente, no conjunto de restrições activas. Em qualquer uma destas regiões a pressão do reactor deve ser mantida no seu limite superior, enquanto que o nível do reactor deve estar no seu limite inferior. Na região do custo mínimo, o caudal de vapor no *stripper* deve estar no seu limite inferior, enquanto que na região próxima da produção máxima deve manter-se no limite superior. Na região da produção máxima pelo menos uma das alimentações deverá ser mantida no limite superior. Os diferentes conjuntos de restrições activas implicam que os graus de liberdade disponíveis para o controlo do processo sejam diferentes consoante a região onde se pretende operar o processo. Grande parte das estruturas de controlo SISO publicadas foram concebidas para operação apenas numa das regiões, considerando um conjunto fixo de restrições activas.

Para além das questões já referidas devem ainda ser referidas as dificuldades numéricas surgidas na obtenção dos resultados apresentados. Estas dificuldades devem-se, sobretudo, à não-linearidade, à instabilidade em ciclo aberto e à existência de variáveis com tempos de resposta com diferentes magnitudes. É expectável que estas características possam condicionar a aplicação de algoritmos de optimização do comportamento dinâmico que utilizem o modelo do processo.

Os resultados obtidos indicam que o controlo adequado do processo não será possível controlando isoladamente cada uma das unidades, devendo ser utilizada uma estrutura de controlo que tenha em consideração a interacção entre as diferentes unidades e a forma como os objectivos de controlo se reflectem nas suas condições de operação. A estrutura de controlo deve ser suficientemente flexível para permitir a operação quer na região do custo mínimo, com qualquer caudal de produto, quer nas condições de produção máxima. Os graus de liberdade disponíveis em cada uma destas condições são diferentes, dependendo também da mistura  $G/H$  pretendida no produto, o que dificulta a selecção dos emparelhamentos de variáveis a utilizar com controladores SISO. A isto acresce a não-linearidade do processo que torna critérios como RGA, ou outros baseados em aproximação lineares, apenas válidos numa região em torno do ponto em que é feita a linearização. Isto mesmo se verifica nos resultados publicados por outros autores em que é sempre necessário o recurso a heurísticas e a testes de simulação para selecção dos emparelhamentos mais adequados.

Todas estas características levam a que, para a solução do problema de controlo do processo TE, seja aconselhada a utilização de controladores baseados em modelos, tais como controladores de carácter predictivo. Estes controladores estabelecem uma perspectiva global (multivariável) do processo, permitindo que as acções de controlo sejam determinadas em função das condições em toda a instalação, e não apenas de cada uma das unidades isoladamente. Estes controladores permitem ainda, de

forma explícita, lidar com restrições quer nas variáveis de saída, quer nas variáveis de entrada, sendo capaz de se adaptar a diferentes combinações de restrições activas. Os resultados obtidos, bem como a vasta gama de condições em que o processo pode ser operado, indicam também que será desejável a utilização de controladores não-lineares. De referir ainda que estes controladores baseados em modelos poderão ser utilizados isoladamente ou, por exemplo, numa estrutura de controlo com duas, ou mais camadas (Qin e Badgwell, 2003; Tatjewski, 2007). Neste caso, a estrutura de nível mais baixo será responsável pela regulação do processo, enquanto que na camada superior é utilizado um controlador também baseado em modelos que, de acordo com as condições do processo, estabelece objectivos (ou referências) para a camada reguladora.

Deve ainda ser referida a importância da existência de um modelo detalhado do comportamento dinâmico do processo, tal como o utilizado neste estudo. Para além de poder ser utilizado por controladores baseados em modelos, permite adquirir uma compreensão do comportamento do processo que dificilmente seria possível obter apenas a partir do funcionamento do processo. Esta compreensão, e a possibilidade de realizar testes de simulação, são fundamentais no projecto de uma estrutura de controlo satisfatória para um problema com esta complexidade.

## Capítulo 4

# Controlo predictivo do processo Tennessee Eastman

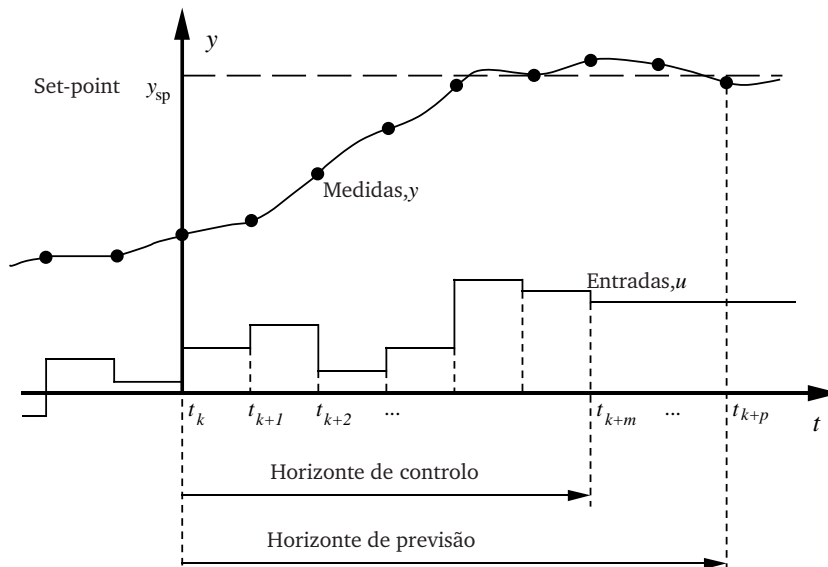
Neste Capítulo são apresentados os resultados obtidos por aplicação de uma estratégia de controlo óptima ao processo Tennessee Eastman (TE). Esta estratégia de controlo aplica a formulação de optimização dinâmica sequencial proposta por [Oliveira \(1994\)](#) para problemas de controlo predictivo não-linear (*nonlinear model predictive control*, NMPC) de processos químicos.

Inicialmente é feita uma descrição da formulação original, bem como das extensões desenvolvidas para aplicação ao processo TE, sendo apenas apresentados os detalhes que se consideram essenciais para a apresentação dos resultados obtidos. Uma descrição mais detalhada pode ser encontrada no trabalho citado. São depois apresentados os resultados de controlo do reactor considerado isoladamente e de todo o processo.

A aplicação desta formulação a um processo com as características do processo TE permite identificar algumas das suas limitações. Neste Capítulo, estas limitações são identificadas e ultrapassadas pela utilização de uma estrutura de controlo em cascata, com a introdução de um compensador interno.

### 4.1 Formulação base de controlo

A formulação de controlo apresentada é uma formulação de controlo predictivo. De forma genérica, neste tipo de formulação pretende-se determinar a trajectória das variáveis manipuladas que corresponde ao valor óptimo previsto de um critério de desempenho, num horizonte temporal futuro. Para que tal seja possível é necessário prever a evolução do processo com o tempo, o que é feito recorrendo a um modelo dinâmico. As características dos controladores predictivos dependem do tipo de mo-



**Figura 4.1** Formulação de MPC com horizontes móveis.

delo utilizado, da sua linearidade ou não, do critério de desempenho considerado e também da forma como, em cada instante, é resolvido problema de optimização. Informação mais extensa sobre as características deste tipo de controladores, o seu desenvolvimento e a sua aplicação pode ser encontrada em [Qin e Badgwell \(1997\)](#), [Camacho e Bordons \(1998\)](#), [Morari e Lee \(1999\)](#), [Qin e Badgwell \(2000\)](#), [Mayne et al. \(2000\)](#), [Kwon e Han \(2005\)](#), [Bock et al. \(2006\)](#), [Tatjewski \(2007\)](#), [Rawlings e Mayne \(2009\)](#), entre outros.

Na formulação utilizada, o horizonte temporal é dividido em intervalos de dimensão constante, igual a  $T$ , designados intervalos de amostragem. Cada intervalo de amostragem é delimitado por dois instantes de tempo,  $t_k$  e  $t_{k+1}$ , onde  $t_k = kT$ ,  $k \in \mathbb{N}_0^+$ , é o instante de amostragem  $k$ , e  $t_{k+1} = t_k + T$ . São considerados horizontes de controlo e previsão com dimensão finita e iguais a  $t_{ih} = mT$  e  $t_{oh} = pT$ , respectivamente, com  $m, p \in \mathbb{N}^+$  e  $p \geq m$ . Entre o fim do horizonte de controlo e o fim do horizonte de previsão as variáveis manipuladas mantêm-se constantes. A implementação deste tipo de estratégia é ilustrada na [Figura 4.1](#).

No instante  $t_k$  o estado actual do processo,  $x(t_k)$ , é fornecido ao controlador. Resolvendo um problema de optimização, é determinada a trajectória das variáveis manipuladas no horizonte de controlo que, partindo deste estado, optimiza o desempenho do processo no horizonte de previsão. Desta trajectória apenas são implementadas as acções de controlo do primeiro intervalo de amostragem,  $[t_k, t_{k+1}[$ . Findo este intervalo de amostragem, é novamente obtido o estado do processo em  $t_{k+1}$ ,  $x(t_{k+1})$ , e o problema é novamente resolvido, agora a partir deste instante de amostragem. Este tipo de implementação é geralmente designado como controlo com horizontes móveis, ou regressivos (*receding horizon*), e confere ao controlo predictivo caracterís-

ticas de controlo por realimentação (Kwon e Han, 2005). A solução obtida em  $t_k$  é geralmente utilizada como aproximação inicial à solução em  $t_{k+1}$ .

São considerados modelos constituídos por DAE's em que é possível a separação das equações a partir das quais são obtidas as variáveis medidas. No caso geral de DAE's na forma implícita, o modelo do processo pode ser representado na forma genérica

$$F(\dot{x}(t), x(t), u(t), w(t), \theta, t) = 0 \quad (4.1a)$$

$$G(x(t), y(t), \theta) = 0. \quad (4.1b)$$

onde  $x \in \mathbb{R}^{n_x}$  são as variáveis de estado,  $u \in \mathbb{R}^v$  são as variáveis manipuláveis,  $w \in \mathbb{R}^{n_w}$  são variáveis de entrada não manipuláveis,  $\theta \in \mathbb{R}^{n_p}$  são parâmetros do modelo, que se consideram invariantes com o tempo, e  $y \in \mathbb{R}^{n_y}$  são as variáveis medidas. As condições iniciais têm a forma:

$$N(t_0, x(t_0)) = 0$$

Supõe-se que estas condições são consistentes e que são satisfeitas as condições necessárias para garantir a existência e unicidade de uma solução, de acordo com o exposto na Secção 2.1.

Como função objectivo do problema de optimização será considerado um critério de desempenho semelhante a (2.5), ou seja o desvio das variáveis medidas e manipuladas relativamente aos seus valores de referência, de forma a manter o processo próximo das condições de operação desejadas. O problema de controlo predictivo a ser resolvido no instante  $t_0$  pode ser escrito como

$$\min_{u(t)} J_2 = \int_{t_0}^{t_0+t_{oh}} (y - y_{sp})^T Q_y(t) (y - y_{sp}) + (u - u_r)^T Q_u(t) (u - u_r) dt \quad (4.2a)$$

$$\text{s.a } 0 = F(\dot{x}(t), x(t), u(t), w(t), \theta, t) \quad (4.2b)$$

$$0 = G(x(t), y(t), \theta) \quad (4.2c)$$

$$0 = N(t_0, x(t_0)) \quad (4.2d)$$

$$u^L \leq u \leq u^U \quad (4.2e)$$

$$y^L \leq y \leq y^U \quad (4.2f)$$

$$x^L \leq x \leq x^U \quad (4.2g)$$

$$u(t) = u(t_0 + t_{ih} - T), \quad t \in [t_0 + t_{ih}, t_k + t_{oh}[, \quad (4.2h)$$

onde  $y_{sp}$  e  $u_r(t)$  são as trajectória de referência para as variáveis medidas e manipuladas, respectivamente. A matriz  $Q_y$  deve ser positiva semi-definida e a matriz  $Q_u$  positiva definida, de forma a assegurar a existência de uma solução. Estas matrizes podem ser entendidas como pesos que atribuem diferentes importâncias relativas aos desvios das variáveis, funcionando como parâmetros de sintonização do controlador,

ajustados de acordo com os objectivos de controlo pretendidos.

#### 4.1.1 Modelo discretizado

Para transformar o problema (4.2) num problema com dimensão finita é necessária a parametrização das variáveis de decisão do problema. Para tal, as variáveis manipuladas serão consideradas constantes e iguais ao seu valor no instante inicial de cada intervalo de amostragem. Assim, no intervalo  $[t_k, t_{k+1}[$ , com  $0 \leq k \leq p - 1$ , as variáveis manipuladas são constantes e iguais a  $u_k = u(t_k)$ . As variáveis de decisão do problema passam a ser  $u_j$  ( $j = 0, \dots, m - 1$ ).

Utilizando esta parametrização, é possível obter uma versão discretizada do modelo do processo, em que as variáveis de estado  $x_{k+1} = x(t_{k+1})$  são a solução em  $t_{k+1}$  das equações do modelo (4.1a). Esta solução é obtida pela integração destas equações no intervalo de amostragem  $[t_k, t_{k+1}[$ :

$$F(\dot{x}(t), x(t), u_k, w(t), \theta, t) = 0, \quad t \in ]t_k, t_{k+1}[ \quad (4.3)$$

No primeiro intervalo de amostragem, as condições iniciais são:

$$N(t_0, x_0) = 0 \quad (4.4)$$

Nos restantes intervalos de amostragem do horizonte, as condições iniciais são:

$$x_k = x(t_k) \quad (4.5)$$

onde  $x(t_k)$  é obtido por integração das equações do modelo no intervalo de amostragem anterior a  $t_k$ ,  $]t_{k-1}, t_k]$ . Admitindo que  $F$  e as condições iniciais são tais que garantem a existência, a unicidade e a diferenciabilidade da solução, podemos definir um operador discreto equivalente a (4.3) e (4.4) ou (4.5), representado por

$$x_{k+1} = \chi(t_k + T; x_k, u_k, w; \theta) \quad (4.6)$$

também continuamente diferenciável relativamente a todos os seus argumentos. As condições que tornam possível a definição deste operador são apresentadas por Oliveira (1994). Admitindo também que  $G$  em (4.1a) é tal que  $y$  é unicamente definido e continuamente diferenciável em ordem a  $x$ , podemos definir um operador discreto equivalente a (4.1b)

$$y_k = v(x_k; \theta)$$

também continuamente diferenciável em ordem a  $x_k$ . Estas equações correspondem a uma versão discretizada do modelo do processo, que constituirão as restrições da versão discreta do problema de optimização.



### Derivadas do modelo discretizado

Para a obtenção da solução do problema de otimização é necessária informação de primeira ordem do modelo do processo. Nomeadamente, são necessárias as primeiras derivadas das variáveis de estado no instante  $t_{k+1}$ ,  $x_{k+1}$ , em ordem às variáveis manipuladas no intervalo de amostragem  $[t_k, t_{k+1}[$ . Estas derivadas são

$$\Gamma_k = \Gamma(T, t_k) = \frac{\partial \chi(t_k + T; x_k, u_k, w; \theta)}{\partial u_k} \quad (4.7)$$

com  $\Gamma_k \in \mathbb{R}^{n_x \times n_u}$ . São também necessárias as primeiras derivadas de  $x_{k+1}$  em ordem às variáveis de estado no início do mesmo intervalo de amostragem (que constituem as condições iniciais para a integração das equações do modelo). Estas derivadas são definidas como

$$\Phi_k = \Phi(T, t_k) = \frac{\partial \chi(t_k + T; x_k, u_k, w; \theta)}{\partial x_k}$$

com  $\Phi_k \in \mathbb{R}^{n_x \times n_x}$ . Ambas as derivadas,  $\Gamma_k$  e  $\Phi_k$ , são obtidas pela resolução das equações de sensibilidades (2.4), simultaneamente com a integração das equações do modelo em cada um dos intervalos de amostragem do horizonte de previsão.

Em cada instante de amostragem  $t_k$  do horizonte de previsão, são ainda necessárias as derivadas de primeira ordem das variáveis de saída,  $y_k$ , relativamente às variáveis de estado no mesmo instante. Estas derivadas são definidas por:

$$C_k = \left. \frac{\partial v(x; \theta)}{\partial x} \right|_{x=x_k} \quad (4.8)$$

Para a obtenção simultânea das variáveis de estado  $x_{k+1}$  e das suas derivadas  $\Phi_k$  e  $\Gamma_k$ , definimos o vector aumentado

$$v = [x_k^T \quad u_k^T]^T$$

com  $v \in \mathbb{R}^{n_p}$  e  $n_p = n_x + n_u$ , contendo todas as variáveis relativamente às quais se pretendem obter derivadas. Definimos também as sensibilidades

$$s_i(t) = \frac{\partial x(t)}{\partial v_i}, \quad i = 1, \dots, n_p$$

em que  $s_i \in \mathbb{R}^{n_x}$  são as sensibilidades dos estados relativamente ao parâmetro  $v_i$ . Recorrendo às equações das sensibilidades, definidas por (2.4), obtemos o sistema de DAE's:

$$\begin{cases} F(\dot{x}(t), x(t), u_k, w(t), \theta, t) = 0 \\ \frac{\partial F}{\partial x} s_i + \frac{\partial F}{\partial \dot{x}} \dot{s}_i + \frac{\partial F}{\partial v_i} = 0, \quad i = 1, \dots, n_p \end{cases}, \quad t \in ]t_k, t_{k+1}[$$

A informação pretendida é obtida por integração destas equações no intervalo de amostragem  $[t_k, t_{k+1}]$ , obtendo-se  $x(t_{k+1})$  e  $s(t_{k+1})$ . As condições iniciais para esta integração são, para o primeiro intervalo de amostragem

$$\begin{cases} N(t_k, x(t_k)) = 0 \\ \frac{\partial N}{\partial x} s_i(t_0) + \frac{\partial N}{\partial v_i} = 0, \quad i = 1, \dots, n_p \end{cases}$$

e, para os restantes intervalos

$$\begin{cases} x_k = x(t_k) \\ s_i(t_k) = e_i, \quad i = 1, \dots, n_x \\ s_i(t_k) = 0, \quad i = n_x + 1, \dots, n_p \end{cases}$$

onde  $e_i$  é o vector unidade de dimensão  $i$ . As matrizes das derivadas de primeira ordem são obtidas a partir das sensibilidades em  $t_{k+1}$

$$\begin{aligned} \Phi_k &= [s_{1,k+1} \quad s_{2,k+1} \quad \dots \quad s_{n_x,k+1}] \\ \Gamma_k &= [s_{n_x+1,k+1} \quad s_{n_s+2,k+1} \quad \dots \quad s_{n_x+n_u,k+1}] \end{aligned}$$

em que  $s_{i,k+1} = s_i(t_{k+1})$ . A solução simultânea das equações do modelo e das sensibilidades pode ser feita utilizando qualquer um dos métodos apresentados na secção 2.4.

As derivadas das variáveis de saída em ordem às variáveis de estado podem ser obtidas a partir de (4.1b), através de

$$\frac{\partial G}{\partial y} \Big|_{\substack{x=x_k \\ y=y_k}} C_{k,i} + \frac{\partial G}{\partial x_i} \Big|_{\substack{x=x_k \\ y=y_k}} = 0, \quad i = 1, \dots, n_x$$

onde  $C_{k,i}$  é a coluna  $i$  de  $C_k$  e  $x_i$  é a variável de estado  $i$ .

### Linearização do modelo

A informação de primeira ordem obtida é utilizada na obtenção de uma aproximação linear do modelo, aproximação esta que constituirá umas das restrições do problema quadrático a resolver em cada iteração do algoritmo de optimização. Esta aproximação é obtida em torno de uma trajectória nominal das variáveis manipuladas:

$$\bar{U} = [\bar{u}_k^T \quad \bar{u}_{k+1}^T \quad \dots \quad \bar{u}_{k+m-1}^T]^T$$

A aplicação desta trajectória ao modelo do processo gera as trajectórias nominais das variáveis medidas e de estado:

$$\begin{aligned}\bar{Y} &= [\bar{y}_k^T \quad \bar{y}_{k+1}^T \quad \cdots \quad \bar{y}_{k+p}^T]^T \\ \bar{X} &= [\bar{x}_k^T \quad \bar{x}_{k+1}^T \quad \cdots \quad \bar{x}_{k+p}^T]^T\end{aligned}$$

Assumindo que (4.6) é continuamente diferenciável, a aproximação de primeira ordem à solução das equações do modelo no instante  $t_{k+1}$  pode ser obtida por

$$\Delta x_{k+1} - \Gamma_k \Delta u_k = \Phi_k \Delta x_k \quad (4.9)$$

onde  $\Delta x_k = x_k - \bar{x}_k$  e  $\Delta u_k = u_k - \bar{u}_k$  são desvios relativamente às trajectórias nominais. Se considerarmos que a matriz do Jacobiano de  $\chi$ ,  $J \in \mathbb{R}^{n_x \times (n_x + n_u)}$ ,

$$J(x_k, u_k) = [\Gamma_k \quad \Phi_k]$$

é contínua no sentido de Lipschitz numa vizinhança  $D \subset \mathbb{R}^{n_x + n_u}$  das trajectórias nominais, então o erro cometido nesta aproximação linear é limitado por (Dennis e Schnabel, 1983):

$$\|\chi(t_k + T, x_k, u_k, w; \theta) - \bar{x}_{k+1} - \Gamma_k \Delta u_k - \Phi_k \Delta x_k\| \leq \frac{\gamma}{2} (\|\Delta u_k\|^2 + \|\Delta x_k\|^2),$$

onde  $\gamma$  é a constante de Lipschitz de  $J$  em  $(\bar{x}_k, \bar{u}_k)$ . De forma semelhante, podemos aproximar as equações das variáveis de saída (4.10) em todo o horizonte de previsão por:

$$\Delta y_{k+1} - C_{k+1} \Delta x_{k+1} = 0 \quad (4.10)$$

Se supusermos que também  $C_{k+1}$  é contínua no sentido de Lipschitz numa vizinhança  $S \subset \mathbb{R}^{n_x}$  da trajectória nominal, o erro cometido nesta aproximação linear é limitado por

$$\|v(x_{k+1}; \theta) - \bar{y}_{k+1} - C_{k+1} \Delta x_{k+1}\| \leq \frac{\gamma}{2} \|\Delta x_{k+1}\|^2$$

onde  $\gamma$  é agora a constante de Lipschitz de  $C_{k+1}$  em  $\bar{x}_{k+1}$ . Nas equações (4.9) e (4.10) as matrizes de sensibilidades  $\Phi_k$ ,  $\Gamma_k$  e  $C_k$  são calculadas durante a integração das equações do modelo com a trajectória nominal  $\bar{U}$ .

As mesmas aproximações podem ser definidas para todos os intervalos de amostragem contidos no horizonte de previsão, resultando em:

$$\begin{cases} \Delta x_{i+1} - \Gamma_i \Delta u_i - \Phi_i \Delta x_i = 0 \\ \Delta y_{i+1} - C_{i+1} \Delta x_{i+1} = 0 \end{cases}, \quad i = 0, \dots, p-1. \quad (4.11)$$

Estas expressões são aproximações lineares do modelo do processo na vizinhança das trajectórias nominais. As expressões obtidas são semelhantes a sistemas lineares

discretos com coeficientes variáveis no tempo (LTV).

### Matriz Dinâmica

De forma a simplificar a notação, serão considerados vectores alargados contendo as trajectórias das variáveis medidas e manipuladas, bem como os respectivos valores de referência. Estes vectores são definidos como

$$U = \begin{bmatrix} u_k^T & u_{k+1}^T & \cdots & u_{k+m-1}^T \end{bmatrix}^T; \quad U_r = \begin{bmatrix} u_{r,k}^T & u_{r,k+1}^T & \cdots & u_{r,k+m-1}^T \end{bmatrix}^T$$

$$Y = \begin{bmatrix} y_k^T & y_{k+1}^T & \cdots & y_{k+p}^T \end{bmatrix}^T; \quad Y_{sp} = \begin{bmatrix} y_{sp,k}^T & y_{sp,k+1}^T & \cdots & y_{sp,k+p}^T \end{bmatrix}^T$$

Partindo da linearização do modelo apresentada, é possível obter uma aproximação de primeira ordem da relação entre as variáveis medidas e as variáveis controladas. Essa aproximação pode ser escrita como

$$\hat{Y} = \bar{Y} + \mathcal{S}_m \Delta U \quad (4.12)$$

$$\mathcal{S}_m = \left. \frac{\partial Y}{\partial U} \right|_{U=\bar{U}} \quad (4.13)$$

onde  $\hat{Y}$  é a aproximação de primeira ordem a  $Y$ ,  $\Delta U \equiv U - \bar{U}$  constitui o desvio relativamente à trajectória nominal, e  $\mathcal{S}_m \in \mathbb{R}^{(n_y p) \times (n_u m)}$  é a matriz dinâmica do sistema. Esta matriz representa as sensibilidades das saídas relativamente às entradas, em torno da trajectória nominal, resultando numa matriz triangular inferior por blocos com a forma

$$\mathcal{S}_m = \begin{bmatrix} S_{11} & 0 & \cdots & 0 & 0 \\ S_{21} & S_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ S_{m,1} & S_{m,2} & \cdots & S_{m,m-1} & S_{m,m} \\ S_{m+1,1} & S_{m+1,2} & \cdots & S_{m+1,m-1} & S_{m+1,m} + S_{m+1,m+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{p,1} & S_{p,2} & \cdots & S_{p,m-1} & S_{p,m} + S_{p,m+1} + \cdots + S_{p,p} \end{bmatrix} \quad (4.14)$$

onde cada elemento pode ser expresso na forma:

$$\mathcal{S}_{m_{ij}} = \begin{cases} 0 & \text{se } j > i, \\ S_{ij} & \text{se } j < m, \\ \sum_{l=j}^i S_{il} & \text{se } i \geq m \text{ e } j = m. \end{cases}$$

As matrizes  $S_{ij}$  podem ser obtidas utilizando a regra da cadeia, i.e., a definição da

derivada de funções compostas, para obter as derivadas de  $y_k$  relativamente a todas as entradas anteriores ao instante  $t_k$ . Por exemplo:

$$\begin{aligned} \text{No instante } t_k: \quad & \frac{\partial y_{k+1}}{\partial u_k} = \frac{\partial y_{k+1}}{\partial x_{k+1}} \frac{\partial x_{k+1}}{\partial u_k} = C_{k+1} \Gamma_k \\ \text{No instante } t_{k+1}: \quad & \frac{\partial y_{k+2}}{\partial u_k} = \frac{\partial y_{k+2}}{\partial x_{k+2}} \frac{\partial x_{k+2}}{\partial x_{k+1}} \frac{\partial x_{k+1}}{\partial u_k} = C_{k+2} \Phi_{k+1} \Gamma_k \\ & \frac{\partial y_{k+2}}{\partial u_{k+1}} = \frac{\partial y_{k+2}}{\partial x_{k+2}} \frac{\partial x_{k+2}}{\partial u_{k+1}} = C_{k+2} \Gamma_{k+1} \end{aligned}$$

Aplicando esta regra a todo o horizonte de previsão obtém-se:

$$S_{ij} = \begin{cases} C_{k+i} \Gamma_{k+j-1} & \text{se } i = j, \\ C_{k+i} \left( \prod_{m=1}^{i-j} \Phi_{k+i-m} \right) \Gamma_{k+j-1} & \text{se } i > j. \end{cases} \quad (4.15)$$

### 4.1.2 Solução do problema de optimização

A função objectivo em (4.2) pode ser aproximada por

$$J_2 = (Y - Y_{sp})^T Q_1 (Y - Y_{sp}) + (U - U_r)^T Q_2 (U - U_r),$$

onde  $Q_1 = \text{diag}(Q_{y_i})$  e  $Q_2 = \text{diag}(Q_{u_i})$ . Os coeficientes destas matrizes podem ser escolhidos de forma a que esta aproximação seja uma aproximação rigorosa do termo integral. No entanto, se o problema for directamente especificado nesta forma, os coeficientes destas matrizes podem ser escolhidos de forma a representarem a importância relativa de cada um dos termos.

Utilizando esta última aproximação e definindo  $E = Y_{sp} - \bar{Y}$  e  $U' = U_r - \bar{U}$ , podemos escrever um problema quadrático que aproxima (4.2) como

$$\min_{\Delta U} J_2 = (E - \mathcal{S}_m \Delta U)^T Q_1 (E - \mathcal{S}_m \Delta U) + (\Delta U - U')^T Q_2 (\Delta U - U') \quad (4.16a)$$

$$\text{s.a. } U^{Ld} \leq \Delta U \leq U^{Ud} \quad (4.16b)$$

$$Y^{Ld} \leq \mathcal{S}_m \Delta U \leq Y^{Ud} \quad (4.16c)$$

onde:

$$\begin{aligned} U^{Ld} &= U^L - \bar{U}, & U^{Ud} &= U^U - \bar{U} \\ Y^{Ld} &= Y^L - \bar{Y}, & Y^{Ud} &= Y^U - \bar{Y} \end{aligned}$$

A trajectória das variáveis de saída,  $\bar{Y}$ , é obtida durante a integração das equações do

modelo utilizando a trajectória nominal das variáveis manipuladas,  $\bar{U}$ . O resultado deste problema são os desvios das variáveis manipuladas relativamente à trajectória nominal.

A função objectivo de (4.16) pode ser reescrita na forma

$$J_2(\Delta U) = c + b^T \Delta U + \Delta U^T H \Delta U$$

onde:

$$c = E^T Q_1 E + U'^T Q_2 U' \quad (4.17a)$$

$$b = -2 \mathcal{S}_m^T Q_1 E - 2 Q_2 U' \quad (4.17b)$$

$$H = \mathcal{S}_m^T Q_1 \mathcal{S}_m + Q_2 \quad (4.17c)$$

A linearização do modelo considerada é equivalente a uma aproximação de Gauss–Newton, com restrições, do problema de controlo não-linear. Como já referido na Secção 2.5.2, esta classe de métodos é adequada a problemas com resíduos pequenos, ou seja, a problemas em que o valor da função objectivo na solução óptima é próxima de zero. Esta situação é representativa da generalidade dos problemas de controlo, especialmente quando são considerados intervalos de amostragem pequenos. Neste caso, a solução inicial, obtida a partir da solução óptima do instante de amostragem anterior, é próxima da solução óptima, pelo que o problema convergirá rapidamente para a solução. Uma das dificuldades associadas à aplicação deste método é o termo da matriz dinâmica,  $\mathcal{S}_m^T Q_1 \mathcal{S}_m$ , em (4.17c), que pode provocar o mau condicionamento numérico da matriz Hessiana. Esta situação ocorre sobretudo em sistemas instáveis e agrava-se com o aumento do tamanho do horizonte predictivo.

Esta formulação de controlo foi implementada computacionalmente por Oliveira (1994). Detalhes sobre esta implementação podem ser encontradas no Apêndice C, onde também são apresentados resultados relativamente à eficiência computacional de diferentes métodos de cálculo das sensibilidades do modelo. Os resultados de controlo apresentados nas Secções 4.2 e 4.3 foram obtidos utilizando as rotinas QPSOL (Gill et al., 1983), para resolução dos problemas quadráticos e DASPK (Li e Petzold, 1999), para resolução das equações do modelo e das sensibilidades.

### 4.1.3 Extensão para atrasos nas saídas

Nem todos os processos químicos podem ser bem descritos por modelos do tipo (4.1a). Estão nesta situação os processos que incluem atrasos devidos, por exemplo ao transporte de material entre unidades e aos equipamentos de medição utilizados. Uma situação frequente nestes processos é a existência de atrasos apenas nas equações das saídas. Os atrasos observados podem dever-se, por exemplo, à utilização de instru-

mentação de medida com tempos de resposta relativamente elevados, que não podem ser desprezados. Um exemplo desta situação é a utilização de analisadores para medir a composição de correntes. Nesta secção serão considerados os processos com atrasos constantes nas saídas, descritos por modelos da forma

$$\dot{x} = f(x, t) \quad (4.18a)$$

$$y = g(x(t - \alpha)) \quad (4.18b)$$

onde  $\alpha \in \mathbb{R}^+$  é constante. Nestas expressões assume-se que  $\alpha < t_{oh}$ . A análise deste tipo de atrasos foi considerada por Oliveira (1994) conjuntamente com atrasos nos estados e nas entradas. No caso em que os atrasos nas saídas são constantes é possível adaptar o formalismo da lei de controlo apresentada, de forma a simplificar a sua aplicação. A existência de atrasos nas variáveis medidas implica alterações apenas na forma como é obtida a informação de primeira ordem do modelo, ou seja, na forma como são utilizadas as sensibilidades calculadas durante a integração com a trajectória nominal das variáveis de entrada.

Para simplificar a análise destes modelos em tempo discreto é conveniente dividir o atraso em parte inteira,  $i_\alpha$ , e parte fraccional,  $\Delta_\alpha$  de acordo com

$$\alpha = (i_\alpha - 1)T + \Delta_\alpha$$

com  $i_\alpha \in \mathbb{N}^+$  e  $\Delta_\alpha \in [0, 1[$ . Quando não há atrasos  $i_\alpha = 1$  e  $\Delta_\alpha = 0$ . Serão apenas considerados atrasos nas saídas múltiplos do intervalo de amostragem

$$\alpha = (i_\alpha - 1)T$$

ou seja, sempre com  $\Delta_\alpha = 0$ .

A existência de atrasos nas saídas torna conveniente uma redefinição das suas sensibilidades relativamente aos estados. Em vez de (4.8), neste caso é utilizado

$$C_{k+1}^\alpha = \frac{\partial y(t_{k+i_\alpha})}{\partial x(t_{k+1})} = \frac{\partial g(x_{k+1})}{\partial x_{k+1}} \quad (4.19)$$

para expressar a sensibilidade das saídas no instante  $t_{k+i_\alpha}$  relativamente aos estados no instante  $t_{k+1}$ . Neste caso, as variáveis de entrada no instante  $t_k$  apenas influenciam as saídas nos instantes  $t_{k+i_\alpha}$  e posteriores, sendo

$$\frac{\partial y_{k+i}}{\partial u_k} = 0, \quad i = 1, \dots, i_\alpha - 1$$

No instante  $t_{k+i_\alpha}$  temos

$$\frac{\partial y_{k+i_\alpha}}{\partial u_k} = \frac{\partial y_{k+i_\alpha}}{\partial x_{k+1}} \frac{\partial x_{k+1}}{\partial u_k} = C_{k+1}^\alpha \Gamma_k$$

com a matriz  $\Gamma_k$  definida por (4.7). Esta definição mantém-se, uma vez que não são considerados atrasos quer nos estados quer nas entradas.

Após o instante  $t_{k+i_\alpha}$  as sensibilidades das saídas relativamente às entradas podem ser obtidas utilizando a regra da cadeia, de modo semelhante à situação em que não existem atrasos. No instante  $t_{k+i_\alpha+1}$  tem-se que

$$\begin{aligned} \frac{\partial y_{k+i_\alpha+1}}{\partial u_{k+1}} &= \frac{\partial y_{k+i_\alpha+1}}{\partial x_{k+2}} \frac{\partial x_{k+2}}{\partial u_{k+1}} = C_{k+2}^\alpha \Gamma_{k+1} \\ \frac{\partial y_{k+i_\alpha+1}}{\partial u_k} &= \frac{\partial y_{k+i_\alpha+1}}{\partial x_{k+2}} \frac{\partial x_{k+2}}{\partial x_{k+1}} \frac{\partial x_{k+1}}{\partial u_k} = C_{k+2}^\alpha \Phi_{k+1} \Gamma_k \end{aligned}$$

Procedendo de forma semelhante para os instantes seguintes e definindo as matrizes

$$\mathcal{S}_{mij}^\alpha = \begin{cases} 0 & \text{se } j > i - i_\alpha + 1 \\ S_{i-i_\alpha+1,j}^\alpha & \text{se } j < m \\ \sum_{l=j}^{i-i_\alpha+1} S_{i-i_\alpha+1,l}^\alpha & \text{se } j \leq i - i_\alpha + 1 \text{ e } j \geq m \end{cases} \quad (4.20a)$$

$$S_{ij}^\alpha = \begin{cases} C_{k+i}^\alpha \Gamma_{k+j-1}^\alpha & \text{se } i = j \\ C_{k+i}^\alpha \Phi_{k+i-1}^\alpha \left( \prod_{m=2}^{i-j} \Phi_{k+i-m} \right) \Gamma_{k+j-1} & \text{se } i > j \end{cases} \quad (4.20b)$$

obtem-se uma matriz dinâmica,  $\mathcal{S}_m^\alpha$ , com a estrutura:

$$\mathcal{S}_m^\alpha = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ S_{11}^\alpha & 0 & \cdots & 0 & 0 \\ S_{21}^\alpha & S_{22}^\alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ S_{m,1}^\alpha & S_{m,2}^\alpha & \cdots & S_{m,m-1}^\alpha & S_{m,m}^\alpha \\ S_{m+1,1}^\alpha & S_{m+1,2}^\alpha & \cdots & S_{m+1,m-1}^\alpha & S_{m+1,m}^\alpha + S_{m+1,m+1}^\alpha \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{p-i_\alpha+1,1}^\alpha & S_{p-i_\alpha+1,2}^\alpha & \cdots & S_{p-i_\alpha+1,m-1}^\alpha & \sum_{j=m}^{p-i_\alpha+1} S_{p-i_\alpha+1,j}^\alpha \end{bmatrix} \quad (4.21)$$

A principal diferença entre esta matriz e a apresentada na equação (4.14) é a existência de  $(i_\alpha - 1)$  linhas de matrizes nulas no topo da matriz. Para que a influência



de todas as entradas possa ser contabilizada pelo algoritmo de controlo, o tamanho máximo do horizonte de controlo deve ser  $p - i_\alpha + 1$ , uma vez que as entradas posteriores ao instante  $t_{p-i_\alpha+1}$  não têm qualquer influência nas saídas localizadas dentro do horizonte predictivo. Quando existem atrasos múltiplos, a equação (4.18b) deve ser generalizada. Considerando, por exemplo, dois atrasos constante  $\alpha_1$  e  $\alpha_2$  tem-se

$$y(t) = g(x(t), x(t - \alpha_1), x(t - \alpha_2))$$

onde  $\alpha_1, \alpha_2 \in \mathbb{R}^+$  e  $\alpha_1 \neq \alpha_2$ . A linearização do modelo em cada iteração do algoritmo de controlo permite escrever neste caso

$$\mathcal{S}_{m,proc} = \mathcal{S}_m + \mathcal{S}_m^{\alpha_1} + \mathcal{S}_m^{\alpha_2} \quad (4.22)$$

onde  $\mathcal{S}_m$  é calculada por (4.14) e  $\mathcal{S}_m^{\alpha_1}, \mathcal{S}_m^{\alpha_2}$  por (4.21). Nesta situação o tamanho máximo do horizonte de controlo deverá ser  $p - \max(i_{\alpha_1}, i_{\alpha_2}) + 1$ .

#### 4.1.4 Acção integral

Para eliminar erros estacionários, provenientes quer de desajustes entre o modelo e o processo, quer de perturbações externas, é conveniente introduzir na lei de controlo de um elemento integrador. Na formulação anterior este elemento integrador pode ser incluído na formulação através de uma modificação simples das equações de estado, do tipo

$$\begin{aligned} z_{k+i} &= z_{k+i-1} + K_I(y_{k+i} - y_{sp,k+i}), & i = 1, \dots, p \\ z_k &= z_0 \end{aligned}$$

com  $z_k \in \mathbb{R}^{n_y}$  e  $K_I \in \mathbb{R}^{n_y \times n_y}$ . A matriz  $K_I$ , designada por ganho integral, é determinante na velocidade de resposta do elemento integrador. Utilizando variáveis desvio  $x' = \hat{x} - \bar{x}$  e  $u' = \hat{u} - \bar{u}$ , a aproximação de primeira ordem a  $z_{k+i}$  é dada por:

$$\begin{aligned} \hat{z}_{k+i} &= \hat{z}_{k+i-1} + K_I(\hat{y}_{k+i} - y_{sp,k+i}) \\ &= \hat{z}_{k+i-1} + K_I(\bar{y}_{k+i} + C_{k+i}(\Phi_{k+i-1}x'_{k+i-1} + \Gamma_{k+i-1}u'_{k+i-1})) - K_I y_{sp,k+i} \end{aligned}$$

Definindo  $z'_{k+i} = \hat{z}_{k+i} - \bar{z}_{k+i}$ , obtem-se:

$$\begin{aligned} \bar{z}_{k+i} &= \bar{z}_{k+i-1} + K_I(\bar{y}_{k+i} - y_{sp,k+i}), & i = 1, \dots, p \\ \bar{z}_k &= z_0 \end{aligned}$$

Pode agora ser definida uma aproximação linear para o conjunto das equações de estado do modelo (4.1a) e do integrador de forma a obtermos as respectivas matrizes

das sensibilidades:

$$\begin{aligned}\tilde{x}'_{k+i} &= \begin{bmatrix} x'_{k+i} \\ z'_{k+i} \end{bmatrix} = \tilde{\Phi}_{k+i-1} \tilde{x}'_{k+i-1} + \tilde{\Gamma}_{k+i-1} u'_{k+i-1} \\ &= \begin{bmatrix} \Phi_{k+i-1} & 0 \\ K_I C_{k+i} \Phi_{k+i-1} & I_{n_y} \end{bmatrix} \begin{bmatrix} x'_{k+i-1} \\ z'_{k+i-1} \end{bmatrix} + \begin{bmatrix} \Gamma_{k+i-1} \\ K_I C_{k+i} \Gamma_{k+i-1} \end{bmatrix} u'_{k+i-1}\end{aligned}\quad (4.23a)$$

$$\tilde{y}'_{k+i} = \tilde{C}_{k+i} \tilde{x}'_{k+i} = \begin{bmatrix} C_{k+i} & 0 \\ 0 & I_{n_y} \end{bmatrix} \tilde{x}'_{k+i}\quad (4.23b)$$

O modelo a utilizar na previsão será

$$\tilde{Y} = \tilde{\bar{Y}} + \tilde{\mathcal{F}}_m \Delta U$$

onde:

$$\begin{aligned}\tilde{Y} &\in \mathbb{R}^{n_y p} \equiv [\tilde{y}_{k+1}^T \quad \tilde{y}_{k+2}^T \quad \cdots \quad \tilde{y}_{k+p}^T]^T \\ \tilde{\bar{Y}} &\in \mathbb{R}^{n_y p} \equiv [\tilde{\bar{y}}_{k+1}^T \quad \tilde{\bar{y}}_{k+2}^T \quad \cdots \quad \tilde{\bar{y}}_{k+p}^T]^T \\ \tilde{y} &= [y_i^T \quad z_i^T]^T \quad \tilde{\bar{y}} = [\bar{y}_i^T \quad \bar{z}_i^T]^T, \quad i = k, \dots, p\end{aligned}$$

A matriz  $\tilde{\mathcal{F}}_m$  é obtida utilizando as equações (4.14) a (4.15), substituindo  $C$ ,  $\Phi$  e  $\Gamma$  pelas matrizes aumentadas  $\tilde{C}$ ,  $\tilde{\Phi}$  e  $\tilde{\Gamma}$ . Para utilizar o algoritmo de controlo descrito é ainda necessário definir

$$\tilde{Q}_y = \text{diag}\{Q_y, I_{n_y}\}, \quad \tilde{Y}_{sp} = [Y_{sp}^T \quad 0]^T$$

e substituir as matrizes pelas suas versões aumentadas correspondentes.

### Acção integral com atrasos nas saídas

A implementação da acção integral num processo com atrasos nas saídas requer algumas modificações nas definições das matrizes aumentadas em (4.23). Também aqui serão apenas considerados atrasos constantes e múltiplos do intervalo de amostragem.

Pela definição apresentada em (4.19), para um processo com um atraso  $\alpha = (1 - i_\alpha)T$  nas saídas:

$$C_{k+i}^\alpha = \frac{\partial y_{k+i+i_\alpha-1}}{\partial x_{k+i}}$$

Substituindo a matriz  $C$  pela matriz  $C^\alpha$  em (4.23a), obtém-se:

$$\begin{aligned} z'_{k+i} &= K_I C_{k+i}^\alpha \Phi_{k+i-1} x'_{k+i-1} + z'_{k+i-1} + K_I C_{k+i}^\alpha \Gamma_{k+i-1} u'_{k+i-1} \\ &= z'_{k+i-1} + K_I C_{k+i}^\alpha (\Phi_{k+i-1} x'_{k+i-1} + \Gamma_{k+i-1} u'_{k+i-1}) \\ &= z'_{k+i-1} + K_I y'_{k+i+i_\alpha-1} \end{aligned}$$

Verifica-se que no instante  $t_{k+i}$  as variáveis do integrador,  $z$ , contabilizam os desvios das saídas no instante  $t_{k+i+i_\alpha-1}$ . Definindo uma matriz  $\tilde{C}^\alpha$  tal que

$$\tilde{y}_{k+i+i_\alpha-1} = \tilde{C}_{k+i}^\alpha \tilde{x}_{k+i} = \begin{bmatrix} C_{k+i}^\alpha & 0 \\ 0 & I_{n_y} \end{bmatrix}$$

onde

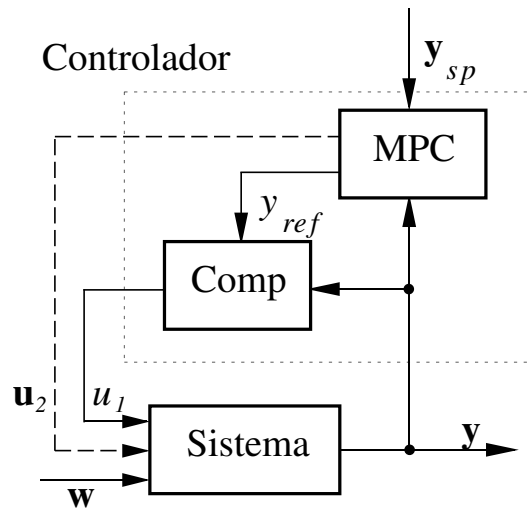
$$\tilde{y}_{k+i+i_\alpha-1} = [y_{k+i+i_\alpha-1}^T \ z_{k+i}^T]^T$$

é possível obter uma aproximação linear para as equações do modelo e do integrador com a mesma estrutura de (4.23).

Conclui-se então, que nesta situação, podem ser utilizadas as matrizes aumentadas definidas em (4.23), apenas substituindo  $C_{k+i}$  por  $C_{k+i}^\alpha$ . A matriz dinâmica é construída utilizando as expressões (4.21) e (4.20). Na situação em que existe mais do que um atraso, por exemplo  $\alpha_1$  e  $\alpha_2$ , constantes, é necessário construir as matrizes aumentadas  $\tilde{C}^{\alpha_1}$  e  $\tilde{C}^{\alpha_2}$ , correspondentes a cada um dos atrasos individuais. A partir destas matrizes, por (4.23a), são obtidas as matrizes  $\tilde{\Phi}^{\alpha_1}$ ,  $\tilde{\Gamma}^{\alpha_1}$ ,  $\tilde{\Phi}^{\alpha_2}$  e  $\tilde{\Gamma}^{\alpha_2}$  com base nas quais é possível obter as matrizes dinâmicas  $\tilde{\mathcal{F}}_m^{\alpha_1}$  e  $\tilde{\mathcal{F}}_m^{\alpha_2}$ , usando (4.21) e (4.20). A matriz dinâmica do processo é finalmente obtida a pela soma destas, como em (4.22).

## 4.2 Controlo do reactor do processo TE

A caracterização do processo TE feita no Capítulo 3 mostra que muitas das características relevantes para o comportamento dinâmico do processo também se encontram no reactor, quando considerado isoladamente. Estas características são a não linearidade, a instabilidade em ciclo aberto, a variação do sinal dos ganhos estacionários e as diferenças de magnitude na velocidade de resposta das variáveis. Para além disso, verifica-se que as variáveis mais críticas do processo são o nível e a pressão do reactor. Por esta razão, antes de considerar a aplicação da formulação de controlo apresentada a todo o processo, esta foi aplicada apenas ao reactor. Os resultados obtidos são aqui apresentados.



**Figura 4.2** Estrutura de controlo em cascata com MPC e compensador.  $y_{ref}$  é a referência do compensador,  $u_1$  é a variável por ele manipulada e  $u_2$  são as variáveis manipuladas pelo MPC.

#### 4.2.1 Estabilização com compensador interno

Devido à instabilidade em ciclo aberto do reactor, verificou-se que o algoritmo de controlo predictivo apresentado na Secção 4.1 não é, por si só, capaz de lidar eficientemente com o problema de controlo desta unidade (Santos et al., 1995). A existência de modos instáveis e de velocidades de resposta muito diferentes levam ao mau condicionamento da matriz dinâmica, o que resulta num problema de optimização numericamente mal condicionado, causando uma importante deterioração do desempenho do algoritmo de controlo. Para atenuar este problema Santos et al. (1996) propõem a utilização de um compensador interno, como alternativa à utilização de formulações de controlo com maior complexidade. Esta abordagem é testada pelos referidos autores no controlo da caixa de entrada de uma máquina de papel.

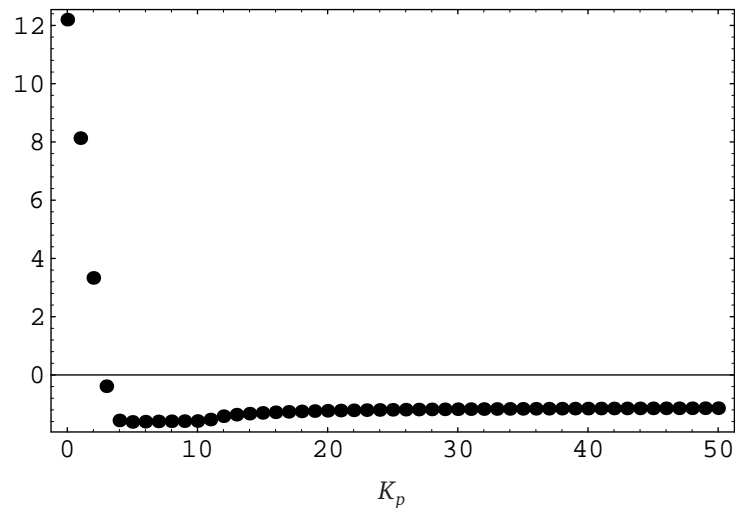
A utilização de um compensador interno foi também considerada para utilização no reactor do processo TE, como forma de ultrapassar os problemas numéricos surgidos na resolução do problema de controlo. A sua utilização resulta numa estrutura de controlo em cascata, como a ilustrada na Figura 4.2.

Nesta estrutura, entre o sistema e o controlador MPC, existem um ou mais compensadores. Estes compensadores são controladores que ajustam directamente algumas das variáveis manipuláveis do processo, com base nos valores de algumas das variáveis medidas e em referências (*set-points*) definidas pelo MPC. Estes controladores têm como principal função estabilizar o processo, não sendo o seu desempenho isolado muito relevante, uma vez que existirá sempre supervisão do MPC. O seu projecto pode, por isso, ser simplificado. Pode ser considerada a existência simultânea de um número arbitrário de compensadores.

Para o controlo do reactor TE, o compensador interno foi seleccionado analisando todos os possíveis emparelhamentos entre variáveis medidas e manipuláveis, utilizando um controlador SISO (*single input, single output*) proporcional. De entre estes, foram seleccionados aqueles que demonstraram ser capazes de estabilizar o reactor. Esta selecção foi feita utilizando o modelo do reactor linearizado em torno das condições de operação correspondentes ao modo base, apresentadas na Tabela 3.7. Com este modelo, foi avaliada a variação, com o ganho proporcional  $K_p$ , do valor próprio da matriz de estado em ciclo fechado com parte real mais elevada:

$$Re(\lambda_{\max}) = \max_{1 \leq i \leq n_s} \{Re(\lambda_i(A + B K_p C))\}$$

A título de exemplo, são apresentados na Figura 4.3 os resultados obtidos para o emparelhamento entre a temperatura do reactor (variável medida) e o caudal da água de arrefecimento (variável manipulável).



**Figura 4.3** Variação da parte real do maior valor próprio,  $\lambda_{\max}$ , com o ganho proporcional,  $K_p$ , para o emparelhamento da temperatura do reactor com o caudal de água de arrefecimento.

Os resultados obtidos mostram a existência de seis emparelhamentos capazes de estabilizar o reactor. Estes emparelhamentos são apresentados na Tabela 4.1.

A implementação dos seis controladores proporcionais, na forma contínua, e a análise da sua resposta às perturbações propostas por [Downs e Vogel \(1993\)](#), permitiu concluir que apenas controlando a temperatura do reactor com o caudal da água de arrefecimento é possível manter o processo dentro dos limites de operação. Assim, o compensador interno seleccionado foi o controlo da temperatura utilizando o caudal da água de arrefecimento, com um ganho proporcional  $K_p = 25$  (normalizado). Nesta estrutura de controlo o MPC deixa de actuar directamente sobre o caudal da água de arrefecimento, para passar a actuar sobre a referência do controlador proporcional.

**Tabela 4.1** Emparelhamentos entre variáveis capazes de estabilizar o modelo linearizado do reactor.

Par	Var. Manipulável	Var. Medida
1	Caudal de água de arref.	Temperatura
2	Caudal de água de arref.	T. saída da água de arref.
3	Velocidade de agitação	T. saída da água de arref.
4	Caudal de D	Temperatura
5	Caudal de água de arref.	Nível
6	Velocidade de agitação	Temperatura

Não há, por isso, variação do número de graus de liberdade do problema de controlo.

Ao nível do problema de optimização, a introdução do compensador reduziu drasticamente o condicionamento da matriz Hessiana dos problemas de optimização correspondentes. Para os mesmos parâmetros de controlo, o número de condição desta matriz passou de valores na ordem dos  $10^6$ , sem compensador, para valores próximos da unidade, com compensador. Esta alteração é suficiente para que o algoritmo de MPC não tenha dificuldades em encontrar uma solução para o problema de controlo. Mesmo nas situações em que, utilizando apenas a formulação MPC original, era possível a obtenção de uma solução de controlo, com a introdução do compensador verifica-se uma redução do número de iterações necessárias para a convergência dos perfis de controlo.

#### 4.2.2 Resultados

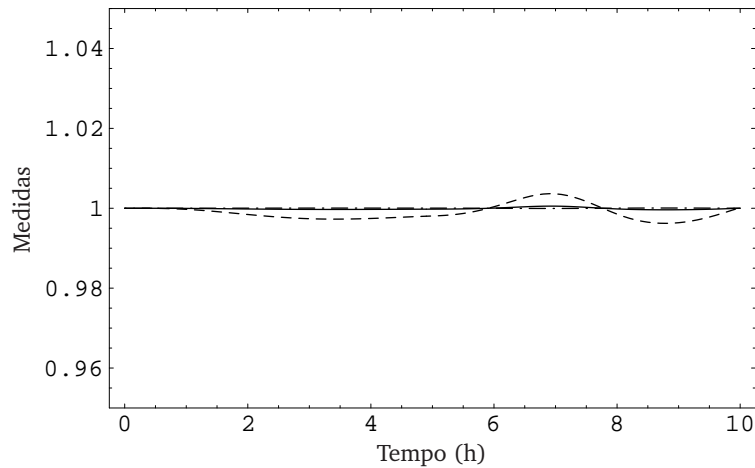
Para avaliar com maior certeza a capacidade da estrutura de controlo proposta lidar eficientemente com o reactor foi necessária a sua implementação e o estudo detalhado do comportamento em ciclo fechado. O compensador interno foi implementado na forma contínua, adicionando a respectiva equação ao modelo. Os resultados apresentados foram obtidos utilizando o mesmo modelo para o controlo e a simulação do reactor, à excepção da informação relativa às perturbações que não é fornecida ao controlador. Foi utilizada a implementação computacional do algoritmo de controlo elaborada por Oliveira (1994), com as extensões descritas nas Secções 4.1.3 e 4.1.4.

O tempo de amostragem utilizado foi 0.1 h, obtido como compromisso entre as diferentes velocidades de resposta das variáveis. Nesta escolha foi também tida em conta a existência de variáveis de saída com atraso.

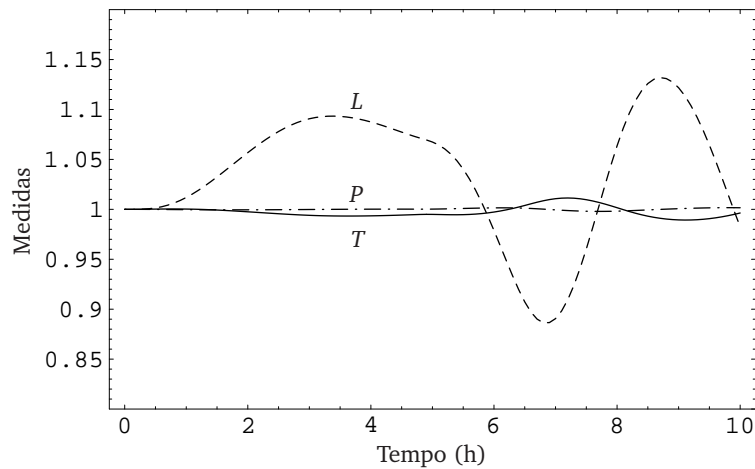
#### Variação aleatória das cinéticas das reacções

Na Figura 4.4 são apresentados os resultados obtidos quando ocorrem variações aleatórias das cinéticas das reacções (perturbação IDV(13) em Downs e Vogel (1993)),

a partir do instante  $t = 0$ . As cinéticas das reacções desviam-se das do modelo do reactor, podendo os desvios atingir os 20%. Em ciclo aberto, esta perturbação causa a paragem do reactor cerca de duas horas após o seu início. Os parâmetros de controlo utilizados foram  $Q_y = I$  e  $Q_u = 10I$ , onde  $I$  representa uma matriz identidade de dimensão adequada. Os horizontes predictivos e de controlo são  $p = 8$  e  $m = 4$ , respectivamente.



(a) Caudais das alimentações de A (— · —), D (---) e E (—).

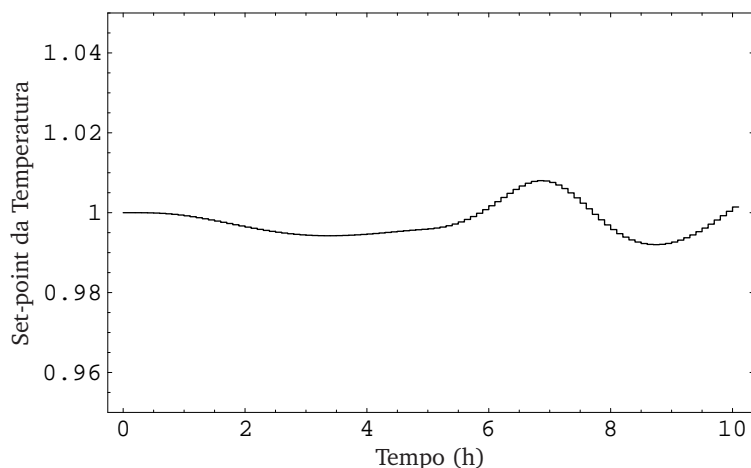


(b) Pressão (P), temperatura (T) e nível (L) do reactor

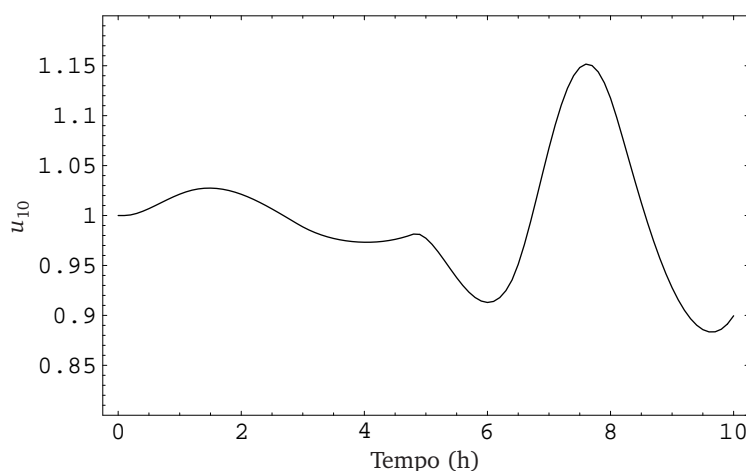
**Figura 4.4** Resposta do reactor em ciclo fechado a variações aleatórias nas cinéticas das reacções.

Para esta perturbação, o controlador consegue manter todas as variáveis medidas próximas dos seus valores de referência, sem acções excessivas sobre as variáveis manipuláveis. As variações dos caudais de alimentação, apresentadas na Figura 4.4a, são muito pequenas e satisfazem as condições impostas.

É de salientar que, apesar de os termos de geração de calor utilizados no modelo serem diferentes dos do reactor, o controlador consegue manter a temperatura do reactor praticamente constante, como se pode observar na Figura 4.4b. Nesta Figura



(c) Set-point da temperatura do reactor



(d) Posição da válvula da água de arrefecimento do reactor

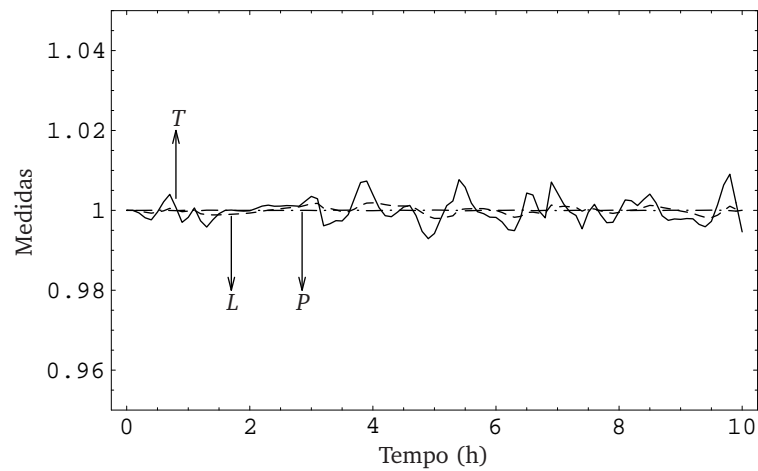
**Figura 4.4 (cont.)** Resposta do reactor em ciclo fechado a variações aleatórias nas cinéticas das reacções.

podemos verificar que o nível apresenta alguma variação mantendo-se, no entanto, afastado dos limites de segurança. A oscilação observada deve-se sobretudo à natureza da perturbação, mais do que ao comportamento do controlador, uma vez que as velocidades de reacção têm grande influência sobre esta variável. Sendo as cinéticas utilizadas pelo controlador diferentes das do reactor é de esperar que verifiquemos algumas dificuldades no controlo do nível. No entanto, são possíveis variações do nível relativamente grandes sem consequências adversas no funcionamento do reactor.

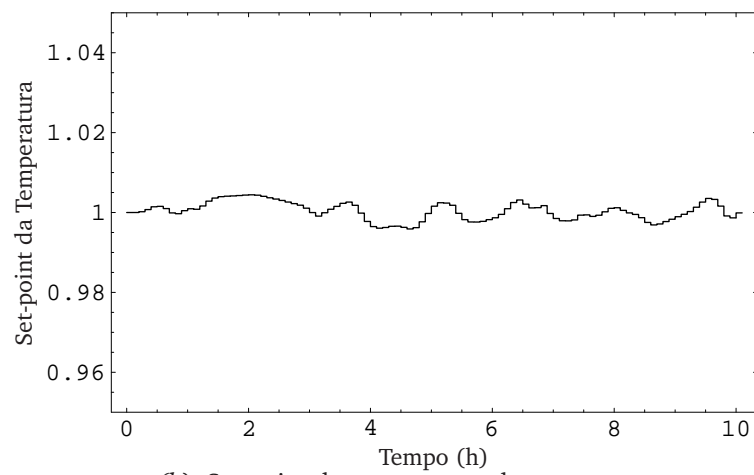
#### **Varição aleatória da temperatura de água de arrefecimento**

Na Figura 4.5 são apresentados os resultados obtidos quando a temperatura de entrada da água de arrefecimento varia aleatoriamente (perturbação IDV(11) em [Downs e Vogel \(1993\)](#)). Os parâmetros de controlo utilizados foram, novamente,  $Q_y = I$ ,  $Q_u = 10I$ ,  $p = 8$  e  $m = 4$ .

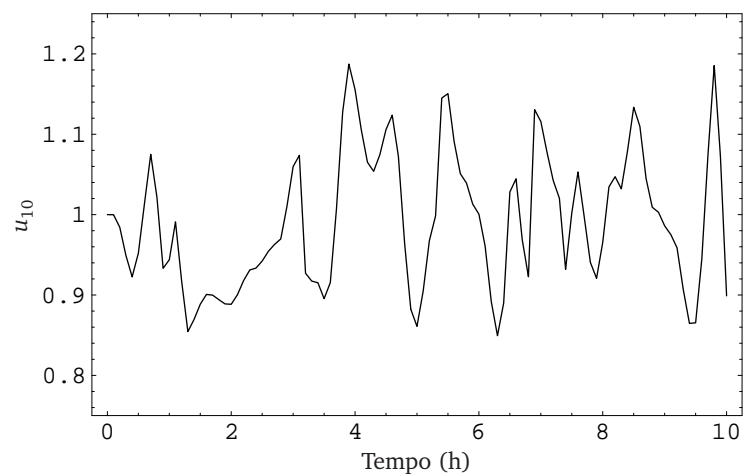




(a) Pressão (P), temperatura (T) e nível (L) do reactor



(b) Set-point da temperatura do reactor



(c) Posição da válvula da água de arrefecimento do reactor

**Figura 4.5** Resposta do reactor em ciclo fechado a variações aleatórias na temperatura de entrada da água de arrefecimento.

Na Figura 4.5a podemos observar que as variáveis medidas mais determinantes no comportamento do reactor mantêm-se sempre muito próximas dos valores de referência, apresentando a temperatura uma variação máxima inferior a 1%. Este comportamento é conseguido manipulando apenas a referência da temperatura do reactor, cuja variação é apresentada na Figura 4.5b, e a válvula do caudal da água de arrefecimento, apresentada na Figura 4.5c. As variações dos caudais das alimentações não são apresentadas por serem muito pequenas podendo ser consideradas constantes.

### 4.3 Controlo do processo TE

Após a obtenção de resultados satisfatórios para o problema de controlo do reactor, o algoritmo foi aplicado ao problema de controlo de todo o processo. Novamente, se verificou que o algoritmo de controlo não é capaz de, por si só, lidar eficientemente com este modelo. As razões apontadas para estas dificuldades são as mesmas já identificadas no controlo do reactor, nomeadamente o mau condicionamento do problema de optimização. Optou-se, por isso, pela utilização de uma estratégia de controlo semelhante à utilizada para o controlo do reactor, ou seja, introduzindo compensador(es) interno(s) entre o MPC e o processo. Nesta secção são apresentados os resultados de controlo obtidos, sendo antes feitas algumas considerações sobre a estrutura de controlo utilizada.

#### 4.3.1 Estrutura de controlo

Os valores próprios do modelo linear do processo em ciclo fechado, com o compensador interno descrito na Secção 4.2.1, mostram que o emparelhamento considerado não é capaz de estabilizar o processo, para qualquer valor do ganho proporcional. Ainda que a estabilização não seja possível apenas com este compensador, e uma vez que se verificou que o reactor é a principal causa da instabilidade do processo, a utilização deste compensador irá, à partida, atenuar as dificuldades no controlo do processo. Isso mesmo se pode verificar na Tabela 4.2, onde são apresentados os números de condição da matriz Hessiana do problema (4.16) para as duas configurações de controlo consideradas (MPC isoladamente e MPC com compensador interno).

Verifica-se que, tal como no problema de controlo do reactor, a introdução do compensador permite reduzir significativamente o número de condição da Hessiana, melhorando o condicionamento do problema de optimização. Estes resultados levaram a que o mesmo compensador utilizado no controlo do reactor isolado fosse também considerado no controlo de todo o processo. Assim, na estrutura de controlo utilizado é mantido o compensador SISO que ajusta o caudal de água de arrefecimento do reactor, utilizando o valor da temperatura do reactor. A referência deste controlador é

**Tabela 4.2** Número de condição da matriz Hessiana para as diferentes configurações de controlo.

	$\kappa(H)$	
	MPC	MPC + compensador
Reactor	$4.6 \times 10^6$	2.3
Processo	$5.6 \times 10^5$	210

estabelecida pelo MPC, que controla todas as restantes variáveis manipuláveis.

Seriam possíveis outras alternativas, como a identificação de um conjunto alternativo de compensadores que permitisse a estabilização do processo. Com este objectivo, a aplicação do procedimento descrito na Secção 4.2.1 embora possível seria extremamente morosa, devido ao grande número de emparelhamentos a considerar. Outra alternativa seria a utilização de qualquer uma das estruturas propostas por outros autores (descritas na Secção 3.2). Esta alternativa não foi considerada por se pretender conseguir o controlo do processo com as menores alterações possíveis à sua estrutura.

### Escolha dos parâmetros de controlo

No problema de controlo do processo TE é necessário ter em consideração a existência de atrasos em algumas variáveis medidas, nomeadamente as composições das correntes. Para evitar formulações mais complexas, optou-se por considerar o atraso na medição da composição do produto (variáveis  $y_{37}$  a  $y_{41}$ ) igual a 0,3 h, e não 0,25 h como no modelo original (Tabela 3.5). Desta forma, todos os atrasos das variáveis medidas serão múltiplos do intervalo de amostragem, permitindo a utilização do procedimento apresentado na Secção 4.1.3.

Considera-se como principal objectivo de controlo a minimização das variações da composição e do caudal do produto em torno das condições de operação do caso base, que constituem a referência. São também penalizados os desvios dos valores das variáveis manipuláveis relativamente aos seus valores neste modo de operação. O intervalo de amostragem utilizado é  $T = 0,1$  h. Os parâmetros de controlo utilizados são  $Q_y = I$  e  $Q_u = 10I$ , para controlo regulador. Os horizontes de previsão e controlo são  $p = 8$  e  $m = 4$ , respectivamente.

Todas as variáveis medidas são fornecidas ao controlador e todas as variáveis manipuláveis são disponibilizadas para o controlo, com excepção do caudal da água de arrefecimento do reactor, ajustado pelo compensador. Não é fornecida à estrutura de controlo qualquer informação sobre as perturbações.

### 4.3.2 Resultados

O desempenho da estrutura de controlo foi testado utilizando as perturbações e mudanças de *set-point* propostas por Downs e Vogel (1993). Os resultados obtidos são apresentados de seguida.

#### Degrau na temperatura da água de arrefecimento do reactor

Nas Figuras 4.6 e 4.7 são apresentados, respectivamente, os perfis das variáveis medidas e manipuláveis para um de grau de  $+5^{\circ}\text{C}$  na temperatura de entrada da água de arrefecimento do reactor no instante  $t = 1$  h. Todas as curvas são normalizadas com os valores das variáveis no modo base de operação.

Os resultados apresentados demonstram a capacidade da estrutura de controlo proposta para rejeitar esta perturbação. Todas as variáveis se mantêm dentro dos limites normais de operação e os desvios observados nas características do produto são muito pequenos.

#### Variação aleatória das cinéticas das reacções

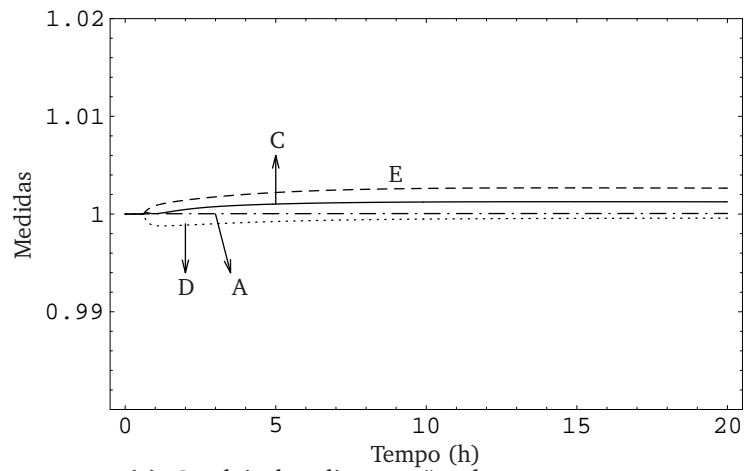
Nas Figuras 4.8 e 4.9 são apresentados os perfis das variáveis medidas e manipuladas para uma variação aleatória das cinéticas das reacções a partir do instante  $t = 1$  h.

Os resultados apresentados permitem concluir que a estrutura de controlo é capaz de lidar com desvios entre o modelo e o processo. De notar que, neste caso, estes desvios se alteram constantemente. Apesar disso, todas as variáveis se mantêm dentro dos limites normais de operação e a variação das características do produto não excede as recomendações de Downs e Vogel (1993).

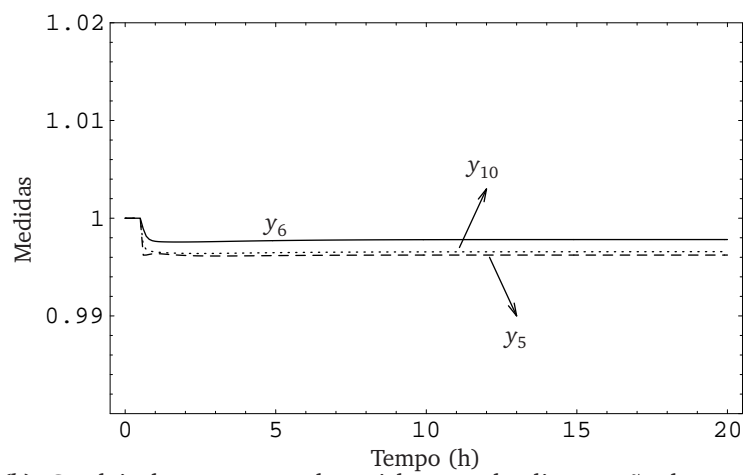
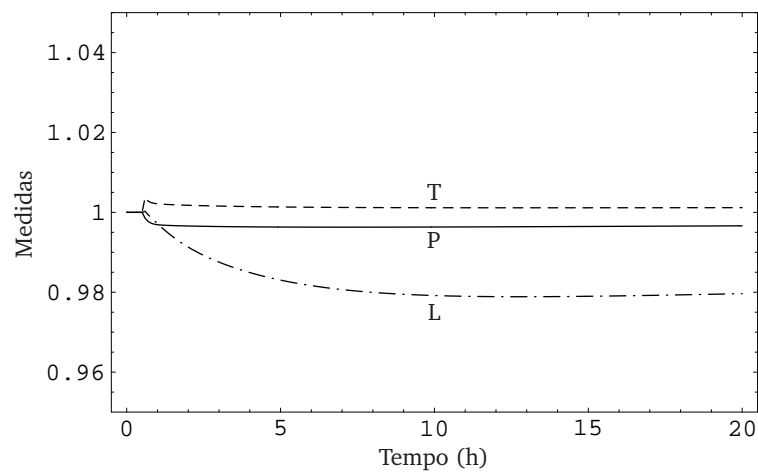
#### Mudança do *set-point* do caudal de produto

Com esta mudança de *set-point* pretende-se diminuir o caudal *mássico* de produto em 15%, de  $14228 \text{ kg h}^{-1}$  para  $12094 \text{ kg h}^{-1}$ , a partir do instante  $t = 0.5$  h. A variável medida cujo *set-point* é alterado é  $y_{17}$ , caudal volumétrico do produto em  $\text{m}^3 \text{ h}^{-1}$  (Tabela 3.3). O seu novo valor foi calculado supondo que a composição do produto é constante, sendo assim possível estabelecer uma relação directa entre o caudal *mássico*, cujo novo valor é conhecido, e o caudal volumétrico, que é medido.

Para que a alteração do caudal de produto seja conseguida de forma eficaz são necessárias alterações nos parâmetros de controlo utilizados, nomeadamente nas matrizes  $Q_u$  e  $Q_y$  e no ganho integral. A matriz das entradas utilizada é  $Q_u = I$ , uma vez que numa mudança de *set-point* não faz sentido penalizar excessivamente as variações das variáveis através das quais a mudança vai ser conseguida. Na matriz dos ganhos das saídas foram também alterados os ganhos correspondentes aos níveis do reactor, do separador e do *stripper*. Para estas variáveis o valor utilizado na linha correspondente

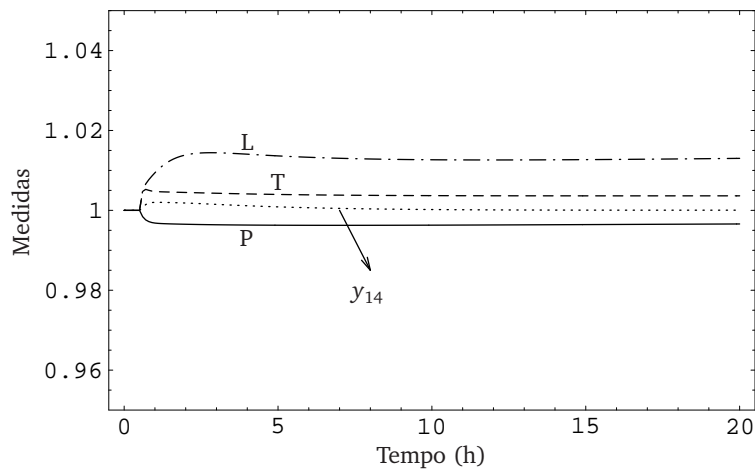
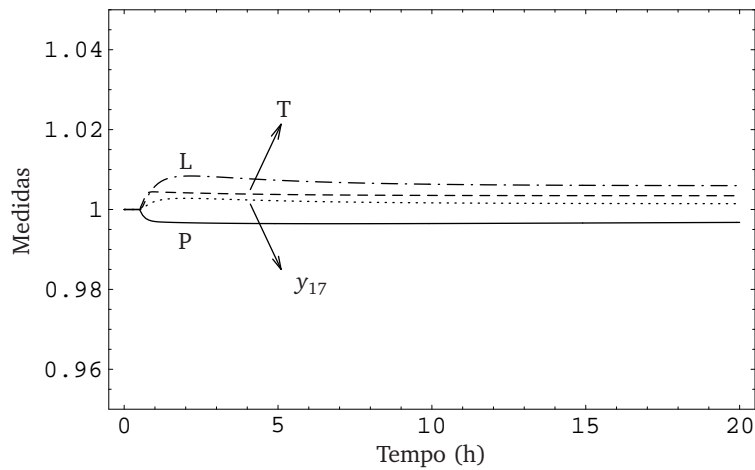
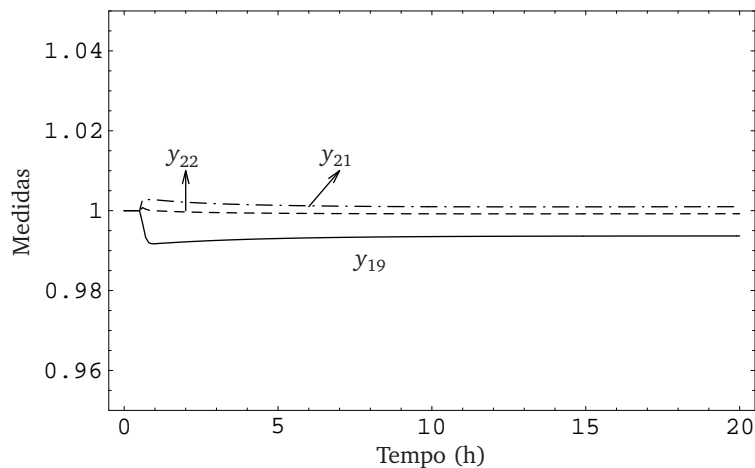


(a) Caudais das alimentações dos reagentes.

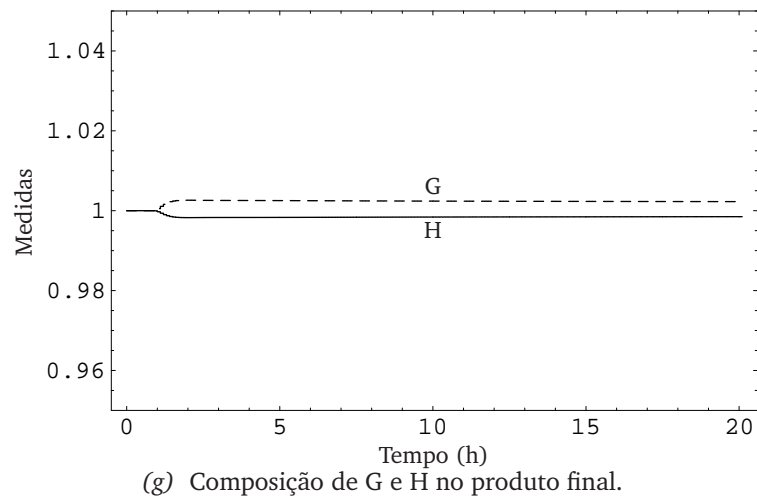
(b) Caudais da purga,  $y_{10}$ , do reciclo,  $y_5$  e da alimentação do reactor,  $y_6$ .

(c) Nível, pressão e temperatura do reactor.

**Figura 4.6** Perfis das variáveis medidas para uma perturbação em degrau de  $+5^\circ\text{C}$  na temperatura de entrada da água de arrefecimento do reactor.

(d) Nível, pressão, temperatura e caudal de base do separador,  $y_{14}$ .(e) Nível, pressão e temperatura do stripper, e caudal de produto,  $y_{17}$ .(f) Caudal de vapor do *stripper*,  $y_{19}$ , temperaturas de saída das águas de arrefecimento do reactor,  $y_{21}$ , e do condensador,  $y_{22}$ .

**Figura 4.6 (cont.)** Perfis das variáveis medidas para uma perturbação em degrau de  $+5^{\circ}\text{C}$  na temperatura de entrada da água de arrefecimento do reactor.



**Figura 4.6 (cont.)** Perfis das variáveis medidas para uma perturbação em degrau de  $+5^{\circ}\text{C}$  na temperatura de entrada da água de arrefecimento do reactor.

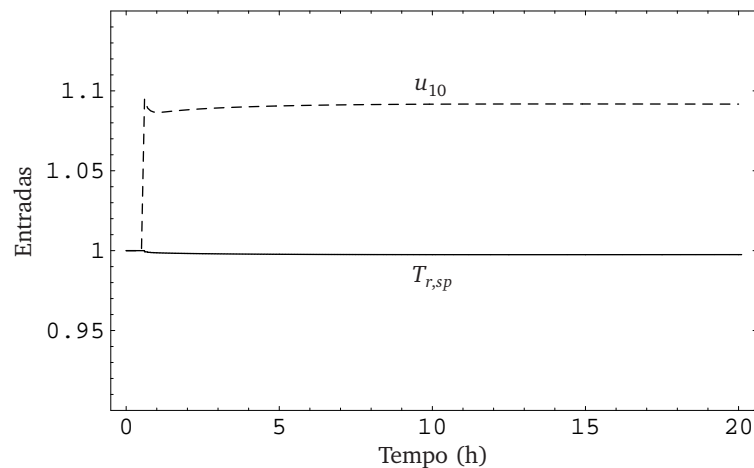
de  $Q_y$  é 10. Com esta alteração pretende evitar-se que a diminuição do caudal de produto seja, nos instantes iniciais, conseguida exclusivamente através da acumulação de líquido naquelas unidades, as únicas onde esta é possível. Este comportamento foi observado nos testes preliminares realizados com o objectivo de sintonizar o sistema de controlo para esta alteração das condições operatórias. É assim evitado que os níveis atinjam valores próximos dos limites máximos, o que constituiria uma dificuldade acrescida no controlo do processo.

Também os ganhos integrais de algumas variáveis de saída foram alterados, nomeadamente do caudal de produto e da composição de G neste, que passaram a ter um ganho  $k_I = 10 \text{ h}^{-1}$ . A alteração do ganho do caudal de produto, variável cujo *set-point* é alterado, é necessária para garantir que esta mudança seja feita rapidamente, e que não exista *offset*. A mudança do ganho da composição tem como objectivo garantir que a composição do produto se mantém constante durante a mudança de caudal. Apenas se aplicou acção integral à composição de G uma vez que no produto apenas existem este composto e H, para além de quantidades residuais de outros compostos. Assim, ao garantirmos que a composição de G se mantém constante garantimos também que a variável que mede a qualidade do produto, a razão mássica  $G/H$  no produto, se mantém constante. Todas as outras variáveis medidas têm ganho integral nulo.

Os resultados obtidos são apresentados nas Figuras 4.10 e 4.11.

Com as alterações nos parâmetros de controlo do MPC a diminuição do caudal de produto é conseguida em menos de 1 hora, sem variações excessivas das variáveis manipuladas, respeitando assim todas as restrições impostas à sua variação, como podemos observar nas Figuras 4.10 e 4.11.

Podemos verificar que o controlador começa a actuar logo no instante inicial de forma



(a) Posição da válvula da água de arrefecimento do reactor e *set-point* da temperatura do reactor.

**Figura 4.7** Perfis de controlo para uma perturbação em degrau de  $+5^{\circ}\text{C}$  na temperatura de entrada da água de arrefecimento do reactor.

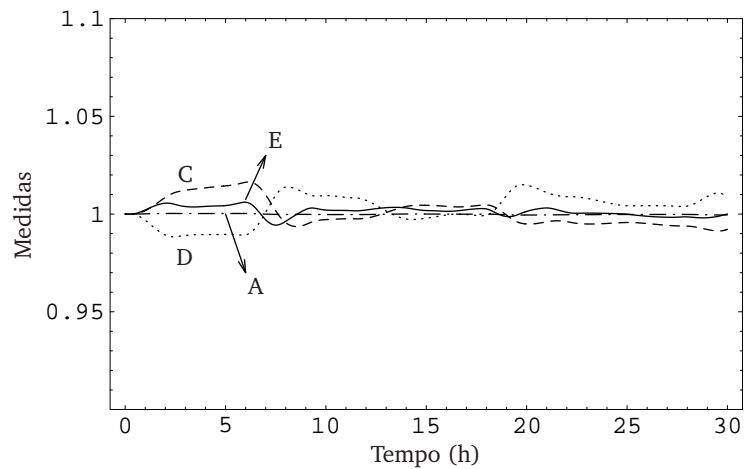
a preparar o sistema para a mudança de *set-point* que ocorre no instante  $t = 0.5$  h. Tal é possível porque, no instante inicial, o horizonte predictivo se estende até  $t = 0.8$  h, ou seja, para além do instante em que está prevista a mudança.

Na Figura 4.10e podemos verificar que a mudança do caudal do produto é conseguida quase instantaneamente, fechando a válvula que controla este caudal. Para que esta diminuição não implique a acumulação de material no processo, também os caudais das alimentações dos reagentes diminuem, o que se pode observar na Figura 4.10a. A diminuição dos caudais dos reagentes não é tão rápida como a do caudal do produto o que tem como consequência um ligeiro aumento dos níveis em todas as unidades, sem que nenhum deles se aproxime dos seus limites.

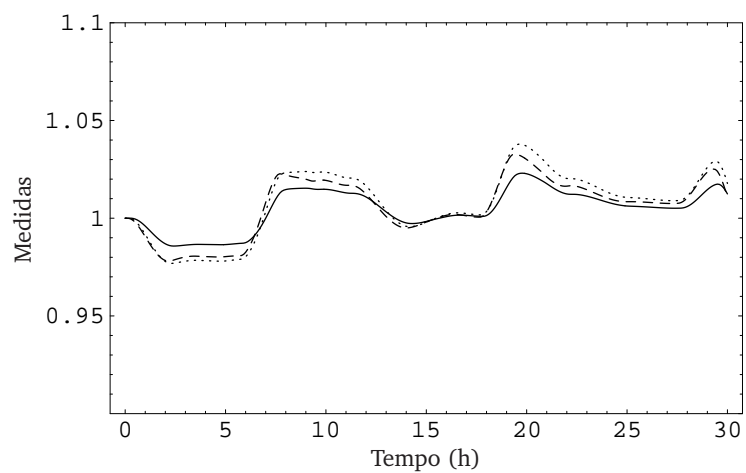
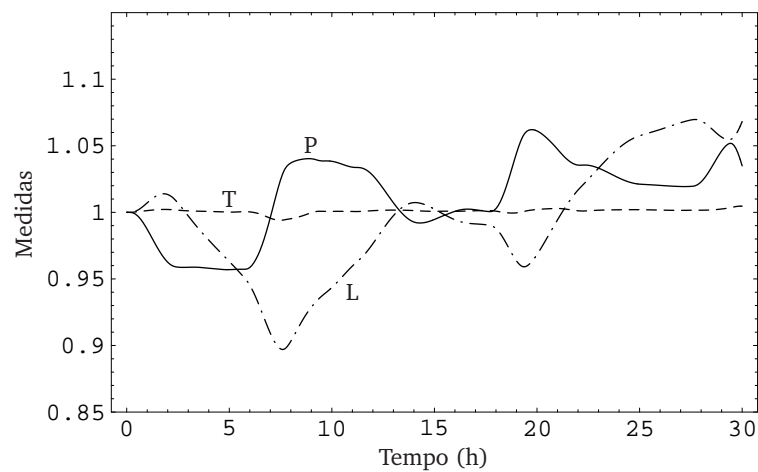
Na Figura 4.10c podemos observar uma diminuição da temperatura do reactor. Esta diminuição é necessária devido à diminuição do caudal de alimentação do reactor, que tem como consequência o aumento do tempo de residência dos reagentes. Se a temperatura fosse mantida constante a relação entre as quantidades de  $G$  e  $H$  produzidas seria alterada, o que implicaria uma degradação da qualidade do produto. Como podemos verificar na Figura 4.11a a diminuição da temperatura do reactor deve-se à acção do controlador e não à diminuição do calor gerado pelas reacções.

A diminuição da temperatura do reactor tem como consequência a diminuição da pressão, apesar de haver um ligeiro aumento do nível. A diminuição da pressão do reactor implica uma diminuição da pressão das restantes unidades. A diminuição da pressão do separador é necessária para que haja transporte de material, uma vez que entre estas unidade este é feito exclusivamente através das diferenças de pressão.



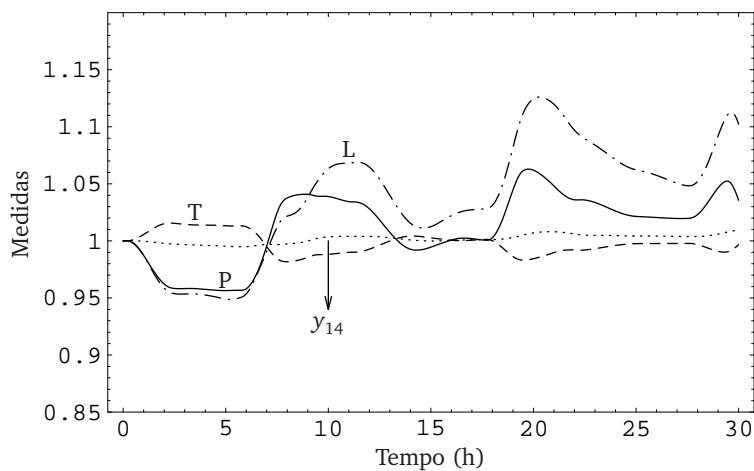


(a) Caudais das alimentações dos reagentes.

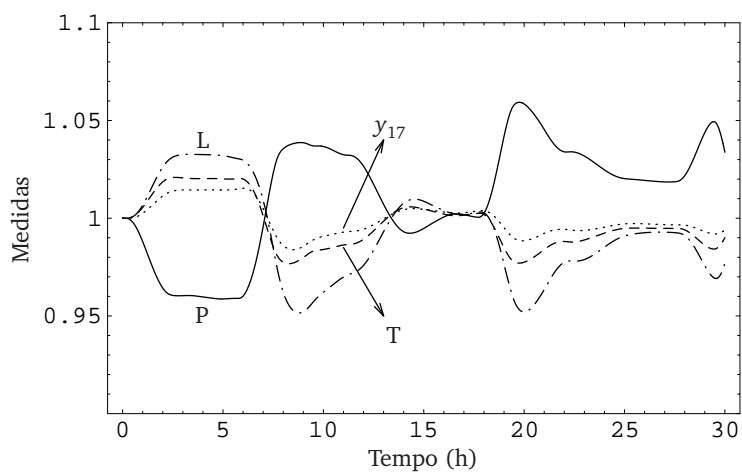
(b) Caudais da purga ( $\cdots$ ), do reciclo ( $--$ ) e da alimentação do reactor ( $—$ ).

(c) Reactor: nível, pressão e temperatura.

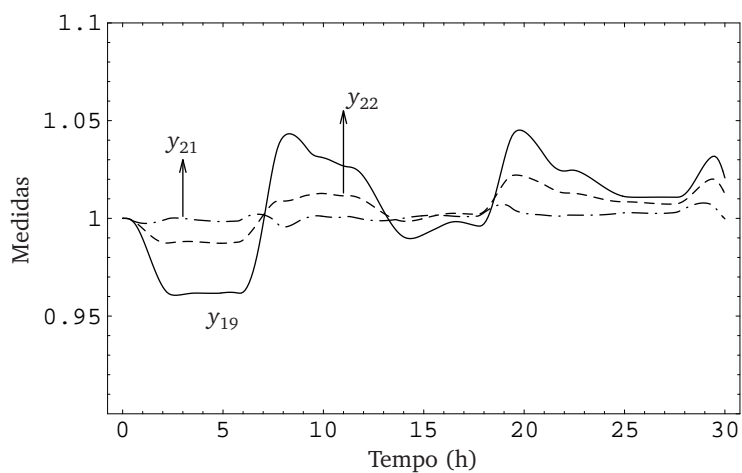
**Figura 4.8** Perfis das variáveis medidas para uma variação aleatória das cinéticas das reacções, perturbação IDV (13) em [Downs e Vogel \(1993\)](#).



(d) Separador: nível, pressão, temperatura e caudal de base,  $y_{14}$ .

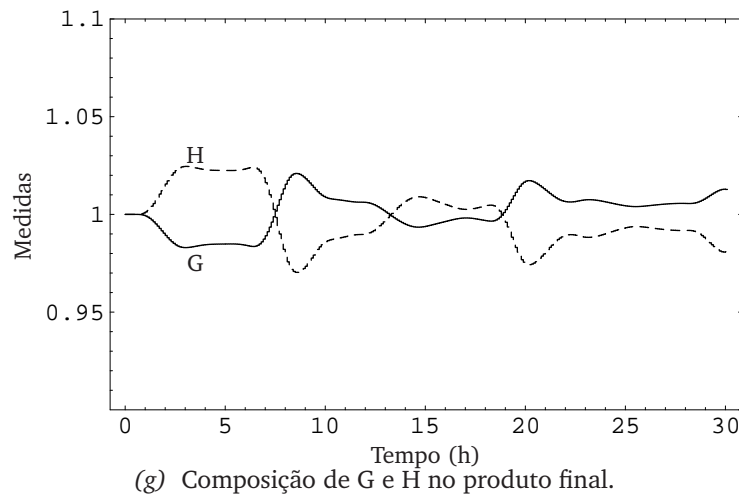


(e) Caudal de produto,  $y_{17}$ , nível, pressão e temperatura do *stripper*.



(f) Caudal de vapor do *stripper*,  $y_{19}$  temperaturas de saída das águas de arrefecimento do reactor,  $y_{21}$ , e do condensador,  $y_{22}$ .

**Figura 4.8 (cont.)** Perfis das variáveis medidas para uma variação aleatória das cinéticas das reacções, perturbação IDV(13) em Downs e Vogel (1993).



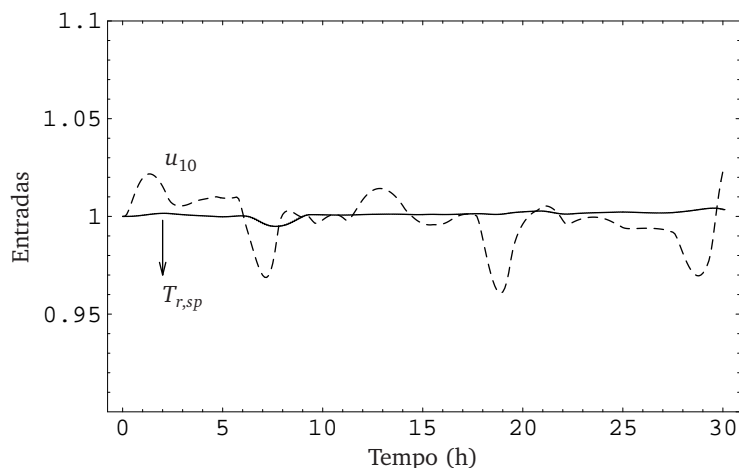
**Figura 4.8 (cont.)** Perfis das variáveis medidas para uma variação aleatória das cinéticas das reacções, perturbação IDV (13) em [Downs e Vogel \(1993\)](#).

#### Mudança do *set-point* da pressão do reactor

O objectivo desta mudança de *set-point* é diminuir a pressão do reactor em 2.2%, de 2705 kPa para 2645 kPa, sem alterações no caudal e na composição do produto. Os parâmetros de controlo utilizados para o conseguir foram  $Q_u = Q_y = I$  e um ganho integral  $k_I = 10 \text{ h}^{-1}$  para a pressão do reactor, o caudal do produto e respectiva composição em G. As razões para estas alterações são semelhantes às apresentadas para a diminuição do caudal de produto.

A diminuição da pressão é conseguida em menos de uma hora. Pode observar-se que esta diminuição é sobretudo conseguida diminuindo o caudal de reciclo, logo o caudal de alimentação do reactor. Um vez que os caudais da purga e do produto se mantêm aproximadamente constantes, um caudal de reciclo mais baixo apenas pode ser conseguido diminuindo a quantidade de material no processo uma vez que não há acumulação no processo. Isto é conseguido diminuindo as alimentações frescas dos reagentes no instantes iniciais, como se pode verificar na Figura 4.12a. Por ocorrer num período de tempo relativamente curto e por não haver oscilações, são satisfeitas as condições requeridas para a variação destes caudais.

Quer o caudal quer a composição do produto final mantêm valores sempre muito próximos da referência, o que é conseguido pela utilização de acção integral. Mais uma vez podemos observar que as pressões do separador e do *stripper* acompanham a variação da pressão do reactor de modo a que continue a existir transporte de material entre elas no sentido desejado. A diminuição da quantidade de material no processo também pode ser observada pela diminuição do nível em todas as unidades. As variações positivas dos níveis do separador e do *stripper* nos instantes iniciais devem-se à necessidade de diminuir rapidamente o nível do reactor de forma a que o volume da fase gasosa aumente. Tal é conseguido aumentando o caudal de saída do reactor que, não havendo um aumento do caudal de saída das outras unidades, implica um



(a) Posição da válvula da água de arrefecimento do reactor e set-point da temperatura do reactor.

**Figura 4.9** Perfis de controlo para uma variação aleatória das cinéticas das reacções, perturbação IDV (13) em Downs e Vogel (1993).

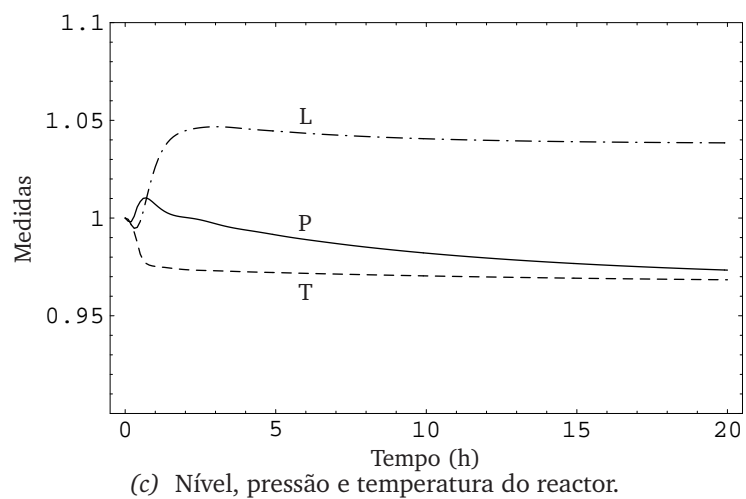
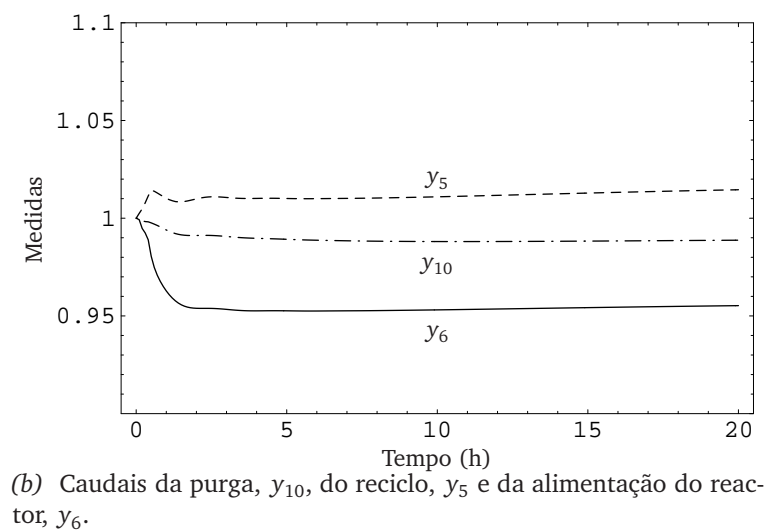
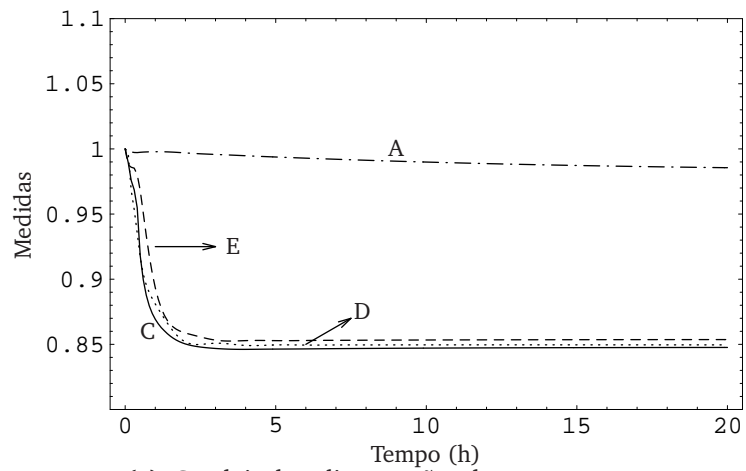
aumento do nível.

Conclui-se que a estrutura de controlo utilizada é capaz de diminuir a pressão do reactor de forma rápida, eficaz e sem *offset*, mantendo sempre o processo longe dos limites de segurança e a qualidade do produto constante, cumprindo assim todos os objectivos de controlo propostos.

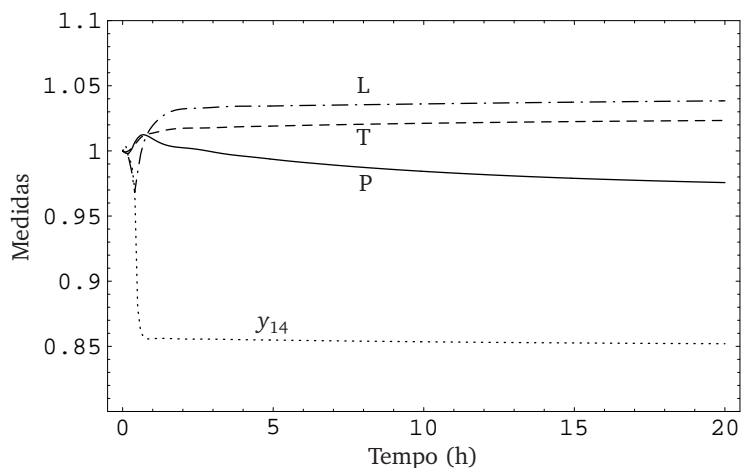
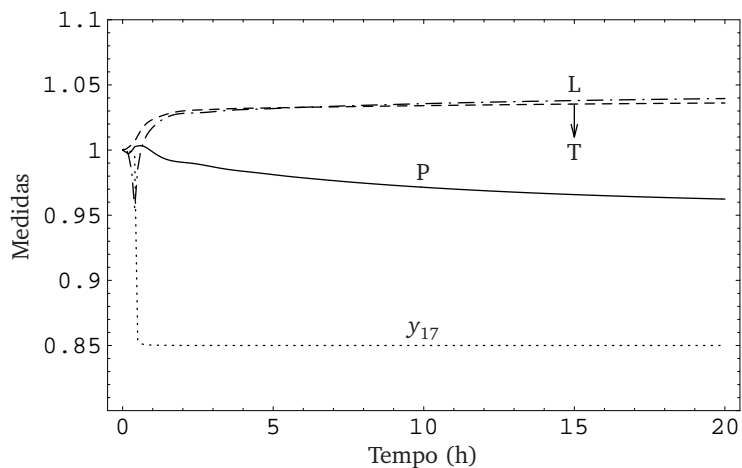
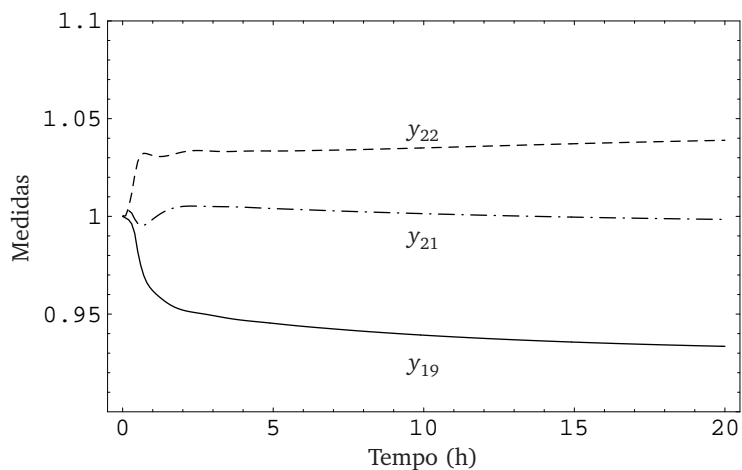
### Mudança do set-point da composição do produto

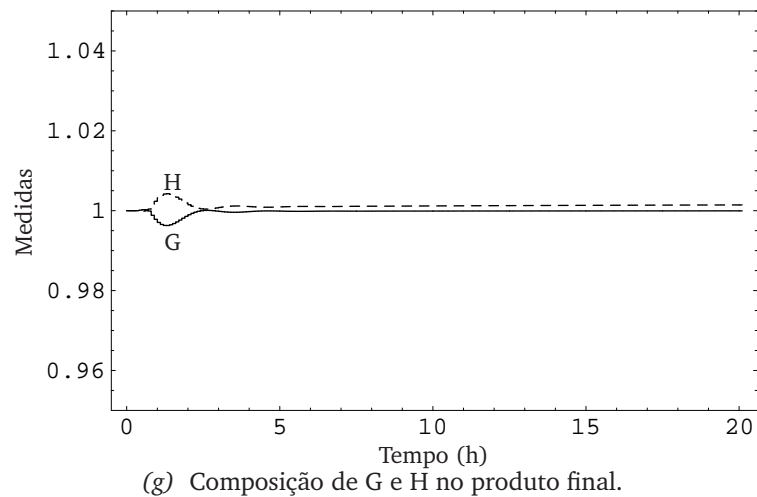
Com a mudança no set-point da composição do produto pretende alterar-se, no instante  $t = 1$  h, os caudais mássicos no produto de 7038 kg/h  $G$  e 7038 kg/h  $H$  para 5630 kg/h  $G$  e 8446 kg/h  $H$ , correspondendo estes valores a uma razão mássica 40 $G$ /60 $H$  no produto final. Uma vez que as variáveis medidas não nos permitem obter directa ou indirectamente esta razão mássica é necessário escolher uma ou mais variáveis medidas que, através de uma mudança dos seus set-points, permitam conseguir a alteração desejada na composição do produto. As variáveis escolhidas para este efeito foram as composições, em %mol, de  $G$  e  $H$  no produto final,  $y_{40}$  e  $y_{41}$  na Tabela 3.5.

Para calcular as composições de  $G$  e  $H$  no produto correspondentes ao novo modo de operação foi necessário considerar que os caudais molares dos restantes compostos se mantêm constantes ao longo da mudança do modo de operação. Fazendo esta aproximação é possível obter valores para as fracções molares de  $G$  e  $H$  no produto que correspondem a uma razão mássica de 40/60. Estas fracções, quando os caudais mássico e molar do produto são iguais aos do caso base, correspondem aos caudais mássicos desejados dos dois compostos. No caso de haver variação de qualquer um dos caudais, as fracções molares obtidas correspondem sempre à razão mássica de-

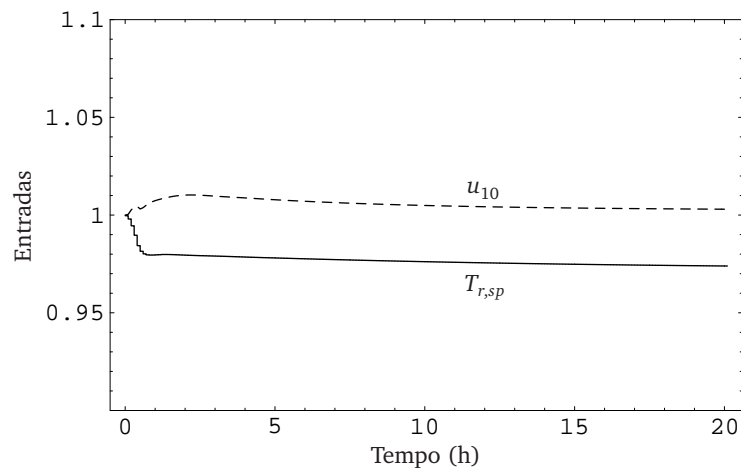


**Figura 4.10** Perfis das variáveis medidas para uma diminuição de 15% do caudal do produto.

(d) Nível, pressão, temperatura e caudal de base do separador,  $y_{14}$ .(e) Nível, pressão e temperatura do stripper, e caudal de produto,  $y_{17}$ .(f) Caudal de vapor do *stripper*,  $y_{19}$ , temperaturas de saída das águas de arrefecimento do reactor,  $y_{21}$ , e do condensador,  $y_{22}$ .**Figura 4.10 (cont.)** Perfis das variáveis medidas para uma diminuição de 15% do caudal do produto.



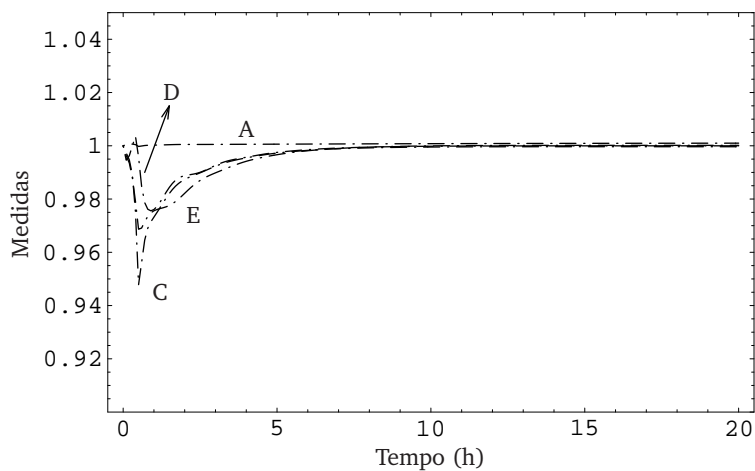
**Figura 4.10 (cont.)** Perfis das variáveis medidas para uma diminuição de 15% do caudal do produto.



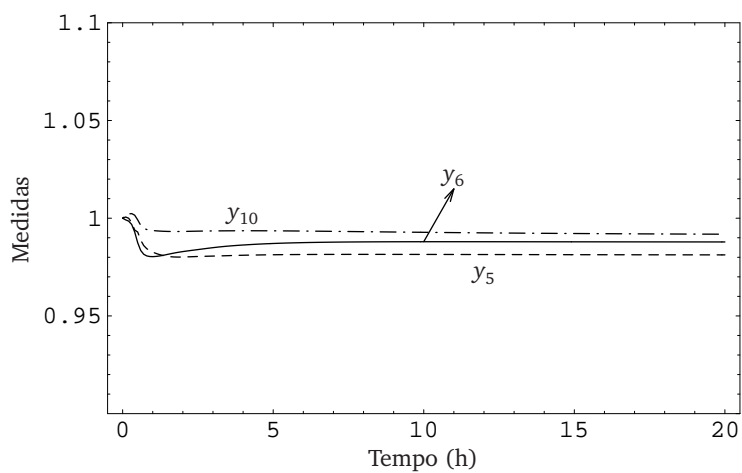
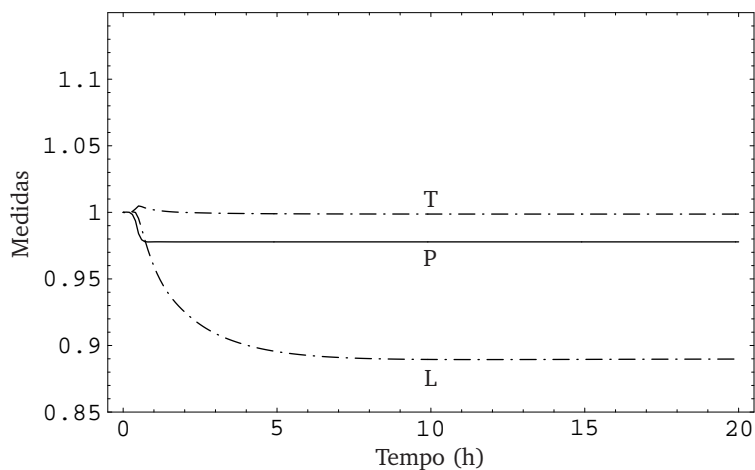
**Figura 4.11** Perfis de controlo para uma diminuição de 15% do caudal do produto.

sejada, ainda que não correspondam aos caudais mássicos pretendidos. Os valores obtidos correspondem a variações de  $-18\%$  e  $+22\%$  dos *set-points* das composições de  $G$  e  $H$ , respectivamente.

Para que a alteração do modo de operação do processo seja conseguida de forma eficaz foram, à semelhança das situações anteriores, necessárias algumas alterações nos parâmetros de controlo utilizados. A penalização atribuída aos desvios das variáveis medidas cujo *set-point* pretendemos alterar é de 10 na matriz  $Q_y$ . Para que a mudança da composição não tenha efeitos adversos no caudal de produto também aos desvios deste é atribuído um peso igual a 10, na matriz  $Q_y$ . Testes preliminares feitos para esta mudança de *set-point* mostraram que uma das formas de esta ser conseguida é através de um aumento da temperatura do reactor. Tal tem como consequência um



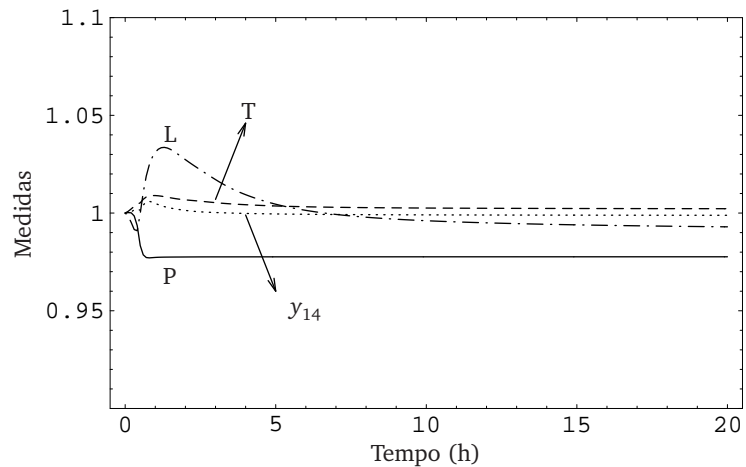
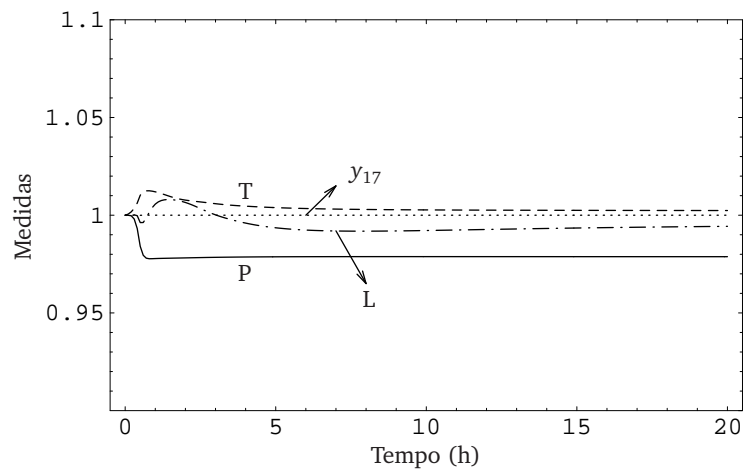
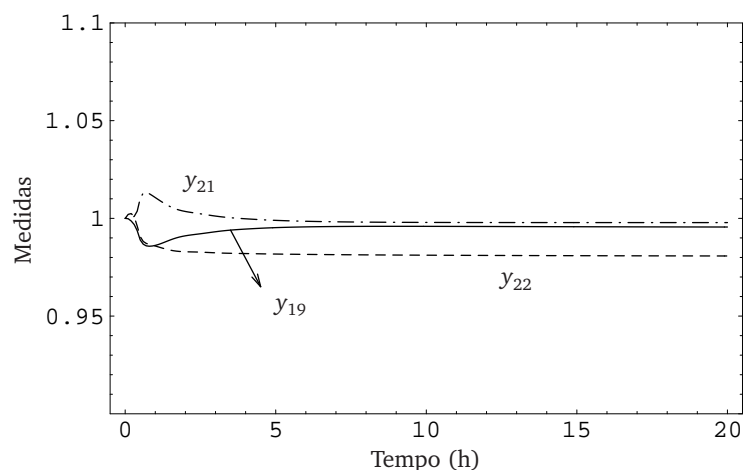
(a) Caudais das alimentações dos reagentes.

(b) Caudais da purga,  $y_{10}$ , do reciclo,  $y_5$  e da alimentação do reactor,  $y_6$ .

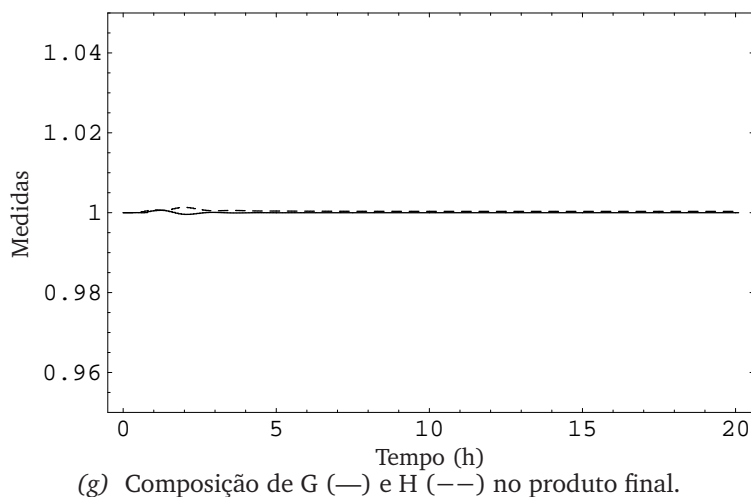
(c) Nível, pressão e temperatura do reactor.

**Figura 4.12** Perfis das variáveis medidas para uma diminuição de 2.2% do *set-point* da pressão do reactor.



(d) Nível, pressão, temperatura e caudal de base do separador,  $y_{14}$ .(e) Nível, pressão e temperatura do stripper, e caudal de produto,  $y_{17}$ .(f) Caudal de vapor do *stripper*,  $y_{19}$ , temperaturas de saída das águas de arrefecimento do reactor,  $y_{21}$ , e do condensador,  $y_{22}$ .

**Figura 4.12 (cont.)** Perfis das variáveis medidas para uma diminuição de 2.2% do *set-point* da pressão do reactor.



**Figura 4.12 (cont.)** Perfis das variáveis medidas para uma diminuição de 2.2% do *set-point* da pressão do reactor.

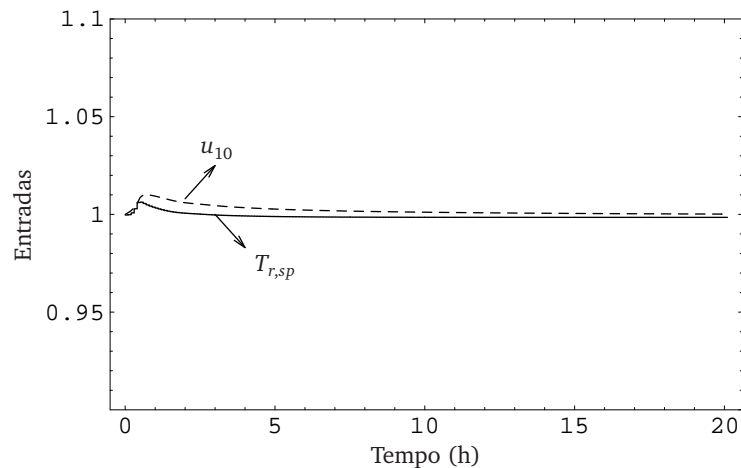
aumento descontrolado da quantidade do subproduto  $F$  acumulada no processo, o que resulta em maiores dificuldades no seu controlo e na diminuição da capacidade de produção. Para evitar esta situação, os desvios nas composições de  $F$  nas correntes de alimentação do reactor e de purga são penalizados com um factor de 10 na matriz  $Q_y$ . Uma vez que é de esperar que para que seja conseguida a mudança de composição do produto sejam necessárias alterações significativas nas condições de operação do processo nenhuma das restantes variáveis medidas do processo é penalizada nos seus desvios. Relativamente às variáveis manipuladas não foi feita qualquer alteração relativamente às restantes mudanças de *set-point*, sendo utilizada uma matriz  $Q_u = I$ .

A aplicação de acção integral às variáveis medidas cujo *set-point* se pretende alterar como forma de evitar *offset* tem que ser ponderada. Uma vez que na corrente do produto apenas existem quantidades residuais de outros compostos que não  $G$  e  $H$ , a variação da quantidade de um deles determina a variação do outro. Por esta razão não é necessário aplicar acção integral a ambas as variáveis que medem as composições, pelo que apenas é aplicada acção integral à composição de  $G$  no produto, com  $k_I = 10$ .

Os resultados obtidos são apresentados nas Figuras 4.14 e 4.15.

Como podemos observar na Figura 4.14g a mudança de composição do produto é conseguida em cerca de uma hora. Para a mesma alteração os melhores resultados apresentados na bibliografia são os conseguidos pela estrutura de controlo proposta por Luyben (1996) em que a mudança da composição do produto é feita em quatro horas.

Nas Figuras 4.14c e 4.15a observa-se que, por acção do controlador, a temperatura do reactor sofre uma variação muito brusca a partir do instante  $t = 0.2$  h, primeiro instante em que a mudança de *set-point* da composição do produto é abrangida pelo horizonte de previsão do controlador. É este aumento da temperatura do reactor, jun-

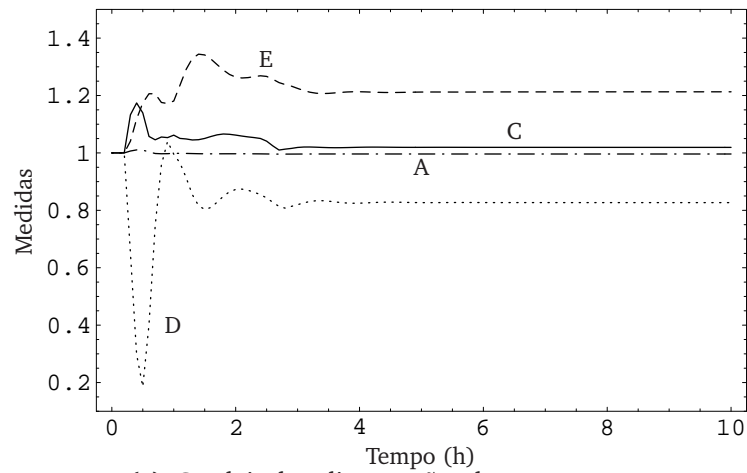


(a) Posição da válvula da água de arrefecimento do reactor e *set-point* da temperatura do reactor.

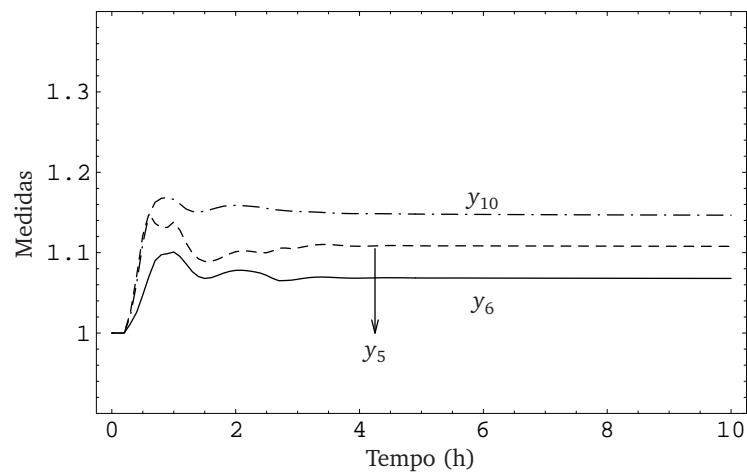
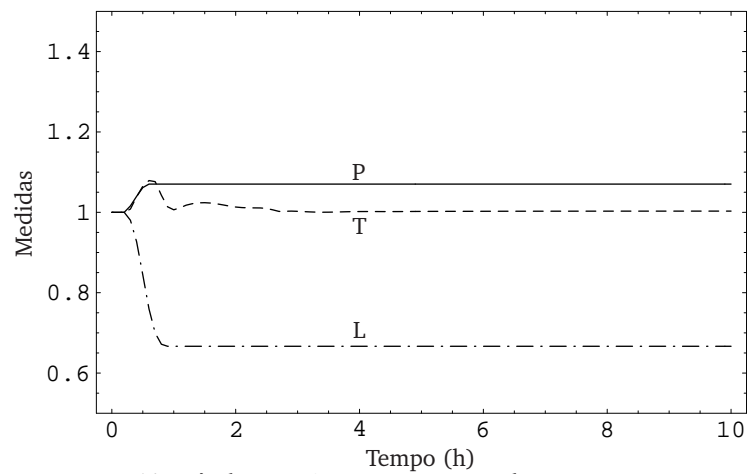
**Figura 4.13** Perfis de controlo para uma diminuição de 2.2% do *set-point* da pressão do reactor.

tamente com as variações do caudais das alimentações frescas de  $D$  e  $E$ , que permite que a mudança da composição do produto seja conseguida tão rapidamente. Pode ainda verificar-se que após algum tempo, e também por acção do controlador, a temperatura do reactor diminui para valores próximos dos do estado inicial. Esta variação é devida ao aumento da produção do subproduto  $F$ , detectada através das medições das composições das correntes de purga e de alimentação do reactor. Com a diminuição da temperatura evita-se o aumento descontrolado da quantidade de  $F$  presente no processo. De forma a que esta variação da temperatura do reactor não tenha efeitos adversos na composição do produto, pode observar-se na Figura 4.15a que, aproximadamente no mesmo instante em que a temperatura do reactor começa a diminuir, o controlador aumenta o caudal de alimentação de  $E$  e diminui o de  $D$ . Tal provoca uma variação da composição destes reagentes no reactor que compensa a perda de selectividade provocada pela diminuição da temperatura. As alimentações dos restantes reagentes mantêm-se aproximadamente constantes uma vez que se pretende que o caudal do produto final se mantenha aproximadamente constante.

Dos dados fornecidos por Downs e Vogel (1993) relativos ao estado estacionário de partida, podemos concluir que a reacção que dá origem a  $H$ , reacção (3.1b), é significativamente mais lenta do que a reacção que dá origem a  $G$ , reacção (3.1a). No estado estacionário de partida o caudal molar da alimentação fresca de  $D$  é maior do que o mesmo caudal de  $E$ , enquanto que na corrente da purga o caudal molar de  $E$  é cerca de dez vezes superior ao de  $D$ . Pode concluir-se que, para as mesmas condições de operação, a reacção (3.1a) é muito mais extensa que a reacção (3.1b), sendo aquela quase completa para as condições de operação do estado estacionário inicial. Tal permite explicar as variações da pressão e do nível do reactor. Na Figura 4.14c observa-se que o nível de líquido no reactor diminui até ao seu limite inferior, dimi-

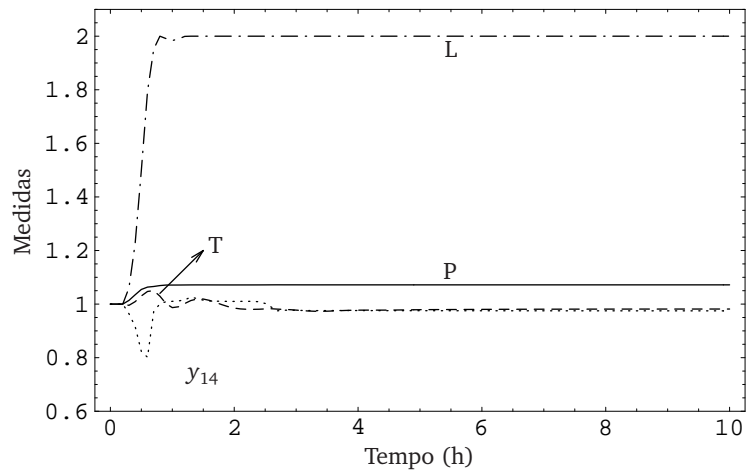
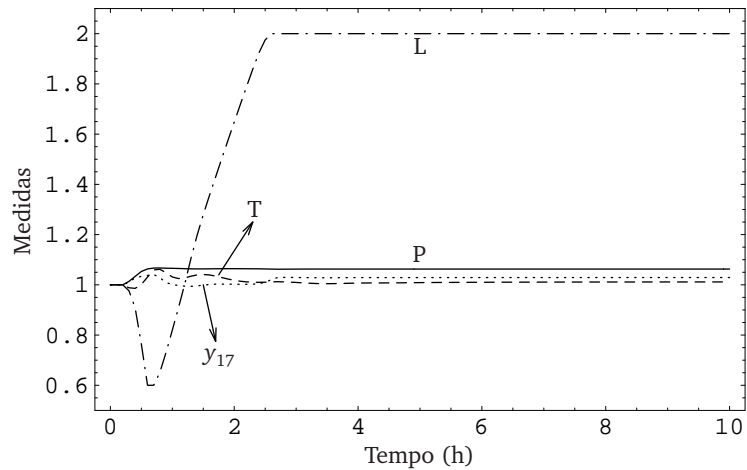
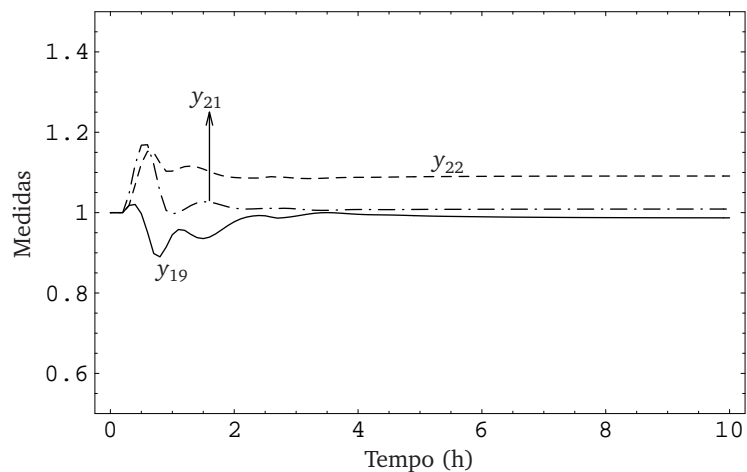


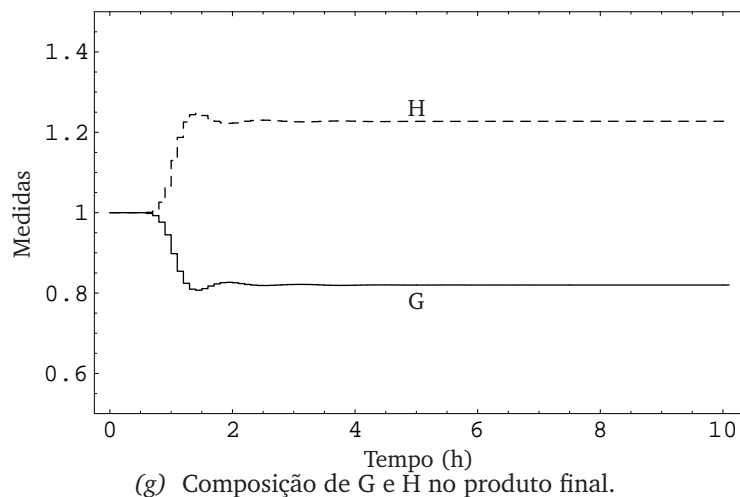
(a) Caudais das alimentações dos reagentes.

(b) Caudais da purga,  $y_{10}$ , do reciclo,  $y_5$  e da alimentação do reactor,  $y_6$ .

(c) Nível, pressão e temperatura do reactor.

**Figura 4.14** Perfis das variáveis medidas para a mudança de *set-point* da composição do produto.

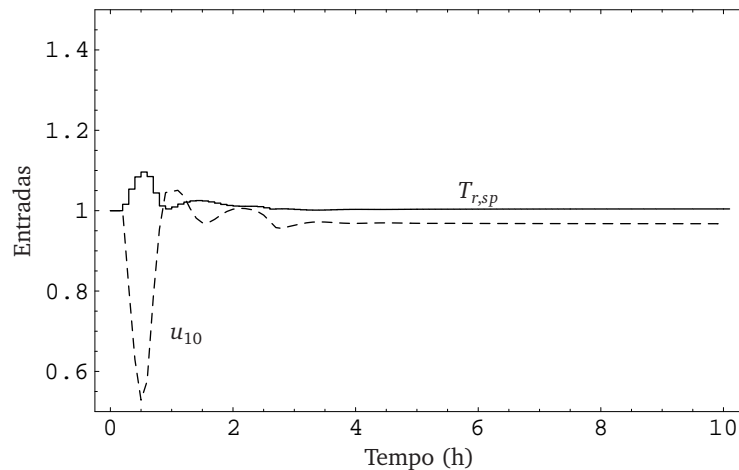
(d) Nível, pressão, temperatura e caudal de base do separador,  $y_{14}$ .(e) Nível, pressão e temperatura do stripper, e caudal de produto,  $y_{17}$ .(f) Caudal de vapor do *stripper*,  $y_{19}$ , temperaturas de saída das águas de arrefecimento do reactor,  $y_{21}$ , e do condensador,  $y_{22}$ .**Figura 4.14 (cont.)** Perfis das variáveis medidas para a mudança de *set-point* da composição do produto.



**Figura 4.14 (cont.)** Perfis das variáveis medidas para a mudança de *set-point* da composição do produto.

nuição esta que é conseguida através de um aumento da vaporização dos compostos líquidos que, por sua vez, é conseguido através do aumento da temperatura do reactor. Esta diminuição tem como consequência um aumento do tempo de residência dos reagentes gasosos no seu interior, logo uma maior extensão das reacções químicas. Uma vez que nas condições de partida a reacção que dá origem a *G* é quase completa, este aumento do tempo de residência favorece sobretudo a produção de *H*. Este aumento da extensão das reacções químicas reflecte-se na diminuição da quantidade dos reagentes *D* e *E* que abandonam o processo pela purga. Na mesma Figura podemos observar que a pressão do reactor aumenta até ao seu limite superior. Uma vez que as velocidades das reacções são proporcionais às pressões parciais dos reagentes, este aumento da pressão compensa a diminuição devida ao aumento do volume de gás, evitando assim a diminuição da capacidade de produção do processo. O aumento da pressão do reactor provoca um aumento semelhante das pressões das restantes unidades, como se pode observar nas Figuras 4.14d e 4.14e.

Nas Figuras 4.14d e 4.14e observa-se um aumento dos níveis do separador e do *stripper* até ao seu limite superior, sendo o aumento no separador mais rápido. Uma vez que é necessário manter o caudal de produto aproximadamente constante, a grande quantidade de material que abandona o reactor durante a diminuição de nível nesta unidade tem que ser armazenada nas unidades a jusante. O grande aumento de nível que se verifica nestas unidades é devido ao facto de a capacidade de armazenagem de líquido no reactor ser cerca de duas vezes superior à capacidade conjunta do separador e do *stripper*. O aumento do nível de líquido no *stripper*, quando este atinge o limite superior, provoca um aumento de 3% no caudal de produto, como se pode observar na Figura 4.14e. Na mesma Figura pode verificar-se que todas as variações do caudal de produto são inferiores ao limite de 5% imposto por Downs e Vogel (1993). O aumento do caudal de reciclo, observado na Figura 4.14b, também se deve



(a) Posição da válvula da água de arrefecimento do reactor e *set-point* da temperatura do reactor.

**Figura 4.15** Perfis de controlo para a mudança de *set-point* da composição do produto.

à necessidade de o processo absorver o material que abandona o reactor.

Na Figura 4.14b podemos verificar que o caudal da purga aumenta. Este aumento deve-se à necessidade de controlar a quantidade de subproduto  $F$  presente no processo. O aumento do caudal da alimentação do reagente  $E$ , através da reacção (3.1c), e o aumento do tempo de residência no reactor favorecem o aumento da quantidade de  $F$  presente no processo. Uma vez que a purga é o único ponto onde este composto pode abandonar o processo, é necessário aumentar o seu caudal de forma a evitar a acumulação excessiva de  $F$ .

Dos resultados apresentados pode concluir-se que a estrutura de controlo utilizada, com parâmetros de controlo adequados, é capaz de mudar a composição do produto em cerca de 1 h, satisfazendo todos os objectivos de controlo propostos. A composição e o caudal do produto, depois de estabilizado o processo, mantêm-se sempre dentro dos limites impostos à sua variação. As correntes das alimentações frescas de  $D$  e  $E$  sofrem alterações significativas, mas estas são indispensáveis uma vez que só mudando as concentrações e as quantidades dos reagentes é possível alterar as quantidades de  $G$  e  $H$  produzidas.

Pode ainda observar-se que quatro variáveis do processo estão nos seus limites. Estas são a pressão do reactor e os níveis do separador e do *stripper* no seu limite superior e o nível do reactor no seu limite mínimo. Tal levanta dificuldades sobretudo ao nível da resolução do problema de optimização, uma vez que é necessário identificar o conjunto de restrições activas. Assim, nos primeiros instantes em que são detetadas restrições activas são necessárias oito iterações do algoritmo de controlo para que seja conseguida convergência. Uma vez que, a partir de determinado momento, o conjunto das restrições activas se mantém inalterado de instante em instante de amostragem, o problema de controlo torna-se mais fácil de resolver, variando entre

três e quatro o número de iterações necessárias para obter convergência.

Para a mesma alteração da composição do produto foram testadas diferentes combinações dos parâmetros de controlo. Nomeadamente foi feita uma simulação utilizando uma matriz  $Q_y = 0$  para todas as variáveis medidas, excepto as composições de  $G$  e  $H$  no produto, para as quais  $Q_y = 10$ . Também foi utilizado  $k_I = 0$  para todas as medidas, excepto a composição de  $G$ , para a qual  $k_I = 10$ . Utilizando estes parâmetros de controlo a nova composição do produto só é conseguida cerca de seis horas após a mudança do *set-point*. Também se verifica que após este instante o controlador não é capaz de estabilizar o processo. As dificuldades de controlo são sobretudo devidas ao aumento descontrolado da quantidade de  $F$  no processo. No instante  $t = 10$  h a composição de  $F$  na corrente de purga é sete vezes superior à do estado estacionário inicial.

Outro problema verificado nesta simulação é o grande número de iterações necessárias para a convergência do perfil de controlo. Uma vez que o processo não estabiliza, o mesmo acontece com o conjunto de restrições activas, pelo que são necessárias mais iterações para o determinar. O número de iterações necessárias para a convergência atinge as trinta e duas. Esta simulação demonstra a importância do controlo da quantidade de  $F$  para um bom desempenho global da estrutura de controlo. Outro facto importante é que a estrutura de controlo utilizada é capaz de manter a pressão do reactor no seu limite superior sem o violar, mesmo quando mais três variáveis estão nos seus limites e o processo ainda não estabilizou

## 4.4 Conclusões

Neste Capítulo foi apresentado o algoritmo de controlo predictivo proposto por [Oliveira \(1994\)](#), para o qual foram desenvolvidas extensões para atrasos nas saídas e aplicação de acção integral. Este algoritmo foi aplicado ao reactor, isoladamente, e ao processo TE. Para que fosse possível a obtenção de resultados, foi necessária a utilização de uma estrutura de controlo em cascata, em que é colocado um compensador interno entre o MPC e o processo. No caso do reactor, este compensador foi escolhido de forma a conseguir a estabilidade antes da aplicação do MPC. Para além de estabilizar o reactor, a introdução deste compensador melhora significativamente o condicionamento numérico do problema de optimização, facilitando a sua resolução. No caso do processo, foi mantida esta estrutura de controlo, ainda que o compensador já não seja suficiente para a estabilização. No entanto, conseguem-se novamente melhorias significativas no condicionamento numérico do problema, o que é suficiente para que a aplicação do MPC tenha sucesso.

Os resultados de controlo apresentados mostram que a estrutura de controlo utilizada, quer no caso do reactor quer no caso do processo, é capaz de rejeitar perturbações e de



conseguir, de forma rápida e eficaz, mudanças de *set-point*, mesmo quando algumas das variáveis se encontram nos seus limites. A introdução de acção integral evita a existência de *offset*.

A utilização de compensadores tem, no entanto, algumas limitações. Uma delas é o facto de a sua escolha de forma sistemática ser um problema extremamente complexo. No caso de processos não-lineares, quando a escolha é feita com base em modelos lineares, acresce ainda o facto de o comportamento dos compensadores apenas ser previsível para condições de operação próximas daquelas em que foi obtido o modelo linearizado. No caso do reactor, a escolha do compensador foi possível, de forma razoavelmente simples, devido ao pequeno número de variáveis, que permite a enumeração de todos os emparelhamentos possíveis entre variáveis medidas e manipuláveis. No caso de processos com maior dimensão esta abordagem já não é viável, sendo necessário recorrer a métodos mais complexos e, muitas vezes, heurísticos. Devem ainda ser consideradas as modificações na estrutura do processo quando são utilizados estes compensadores, uma vez que a solução de controlo obtida é óptima já não para o processo, mas para o sistema constituído pelo processo e pelos compensadores. Esta questão é particularmente relevante quando os compensadores alteram significativamente a dinâmica do processo. No entanto, quando a escolha do compensador é simples, a abordagem apresentada não deixa de ser uma solução interessante, pela sua simplicidade, exequibilidade, e por permitir resolver o problema.



## Capítulo 5

# Reformulação do algoritmo de controlo predictivo

A aplicação da formulação de controlo predictivo apresentada no Capítulo 4 demonstra algumas das limitações apontadas aos métodos sequenciais, nomeadamente as decorrentes do mau condicionamento numérico do problema de optimização e a dificuldade em tratar adequadamente processos instáveis em ciclo aberto. No caso do reactor e do processo TE, estes problemas foram ultrapassados com recurso a alterações na estrutura de controlo, nomeadamente a utilização de um compensador interno entre o processo e o controlador MPC. Esta alteração tornou possível a aplicação do algoritmo de controlo predictivo e a obtenção de soluções de controlo para ambos os sistemas.

Como alternativa à introdução de uma camada adicional de controlo (que pode ser vista como uma alteração ao processo sob controlo predictivo), serão consideradas neste Capítulo modificações na formulação de controlo original que, de forma sistemática, pretendem possibilitar a aplicação directa de métodos sequenciais a uma classe mais abrangente de processos, sem necessidade de alterações na estrutura do processo, mantendo as características desejáveis desta classe de métodos, como a pequena dimensão dos problema de optimização.

Inicialmente serão analisadas as alterações na estrutura de controlo que permitiram a aplicação do algoritmo de controlo predictivo ao processo TE, de forma a perceber quais as consequências destas alterações ao nível das equações e variáveis do modelo. Partindo desta análise, é sugerida uma reformulação do algoritmo de controlo predictivo, ainda baseada em métodos sequenciais, para tentar lidar de forma adequada e sistemática com problemas mal condicionados numericamente. Esta reformulação é complementada com uma estratégia que considera a utilização, quando necessário, de horizontes de previsão com diferentes dimensões. São depois discutidas as vantagens e limitações da abordagem resultante relativamente a outros métodos directos

para a resolução de problemas de optimização dinâmica. Finalmente, são apresentados os resultados obtidos pela aplicação desta abordagem a dois sistemas LTI e ao processo TE.

## 5.1 Motivação

Como mencionado atrás, uma das dificuldades frequentes na resolução de problemas de optimização dinâmica por métodos sequenciais é o mau condicionamento numérico do problema quadrático que é necessário resolver em cada iteração. Esta dificuldade surgiu na aplicação do algoritmo de controlo descrito na Secção 4.1 ao reactor e ao processo TE, impossibilitando a solução do problema de controlo pela sua aplicação directa. No caso deste algoritmo o mau condicionamento é introduzido pela utilização da matriz dinâmica, que representa a linearização das equações do modelo propagadas, em ciclo aberto, ao longo de todo o horizonte de previsão. Através da relação (4.12), esta matriz permite que o problema de optimização seja resolvido utilizando como variáveis de decisão exclusivamente as variáveis manipuladas, que assim correspondem às variáveis independentes do problema de optimização correspondente. As variáveis de estado e medidas são variáveis dependentes, determinadas pela solução das equações do modelo.

Para obtenção dos resultados de controlo apresentados no Capítulo 4 foi necessária a introdução de um compensador, como descrito nas Secções 4.2.1 e 4.3.1. No caso do reactor, este compensador permitiu a sua estabilização, antes da aplicação do MPC. No caso processo TE já não se verificou a estabilização, mas a melhoria conseguida no condicionamento numérico foi suficiente para a resolução, com sucesso, do problema de controlo. Para que as melhorias conseguidas pela utilização do compensador possam ser alcançadas de forma sistemática, é útil identificar que aspectos são responsáveis pela modificação do comportamento do problema. Nomeadamente, é útil compreender quais as alterações na estrutura do problema resultantes da utilização do compensador.

Ao nível do problema de controlo, pode ser estabelecida uma analogia entre a utilização do compensador SISO e a troca de uma variável independente (manipulada) por uma dependente (controlada). No caso do compensador utilizado, a variável manipulável  $u_{10}$ , correspondente à posição da válvula da água de arrefecimento do reactor, deixa de ser uma variável independente, uma vez que o seu valor passa a ser determinado pela temperatura do reactor e pelo valor de referência para esta temperatura, estabelecido pelo MPC. Por outro lado, o valor pretendido para a temperatura do reactor passa a ser livremente estabelecida pelo MPC, através do seu valor de referência, tal como ilustrado na Figura 4.2. Nesta situação, a temperatura do reactor pode ser considerada uma variável independente do problema. Este tipo de altera-

ções nas variáveis do problema foi anteriormente utilizada por [Ricker e Lee \(1995b\)](#) para eliminar os balanços energéticos do modelo do processo TE. No modelo apresentado por estes autores, as temperaturas do reactor e do separador são variáveis de entrada (independentes), usadas para determinar as posições das válvulas de arrefecimento destas unidades (variáveis dependentes) através de controladores SISO, que são considerados parte do processo.

Resulta daqui uma possível ideia base para a reformulação do algoritmo de controlo, que poderá passar pela escolha, de forma sistemática, dos conjuntos de variáveis dependentes e independentes no problema de controlo. Este procedimento deve ser orientado no sentido de conseguir melhorias no condicionamento numérico do problema resultante, devendo, ser possível, ser tornada compatível a sua utilização nos algoritmos sequenciais de controlo predictivo já existentes.

### 5.1.1 Exemplos de teste

Para testar as diferentes possibilidades, nomeadamente, para avaliar eventuais melhorias relativamente à formulação anterior, e interpretar a forma como essas melhorias são conseguidas, a utilização do processo TE não é particularmente adequada. Devido à sua dimensão e complexidade, será mais difícil, com a sua utilização, isolar os aspectos críticos das alterações introduzidas.

Assim, optou-se pela utilização, como exemplos de demonstração, de dois sistemas lineares invariantes com o tempo (*linear time invariant*, LTI), apenas com uma entrada e uma saída. Serão considerados um sistema estável e um sistema instável descritos, respectivamente, pelas funções transferência:

$$\mathcal{G}(s) = \frac{1}{(s+1)(s+2)(s+3)}$$

$$\mathcal{G}(s) = \frac{1}{(s+1)(s-2)(s+3)}$$

Para utilização no algoritmo de MPC, estes sistemas foram discretizados (com um intervalo  $T = 0.5$ ), sendo obtida a sua representação no espaço de estados. Estas representações são, para o sistema estável

$$x_{k+1} = \begin{bmatrix} 0.939 & 0.379 & -0.256 \\ -0.282 & 0.423 & 0.098 \\ -0.587 & -1.357 & -0.164 \end{bmatrix} x_k + \begin{bmatrix} -0.187 \\ 0.047 \\ 0.098 \end{bmatrix} u_k \quad (5.1a)$$

$$y_k = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x_k \quad (5.1b)$$

**Tabela 5.1** Condição numérica da matriz dinâmica e da matriz Hessiana do problema (4.16) para os sistemas (5.1) e (5.2).

Sistema	$\kappa(\mathcal{S}_m)$	$\kappa(H)$
LTI estável	30.1	8.39
LTI instável	<i>singular</i>	<i>singular</i>

e, para o sistema instável,

$$x_{k+1} = \begin{bmatrix} 1.105 & 0.601 & 0.102 \\ 0.615 & 1.618 & 0.396 \\ 2.379 & 2.598 & 0.825 \end{bmatrix} x_k + \begin{bmatrix} 0.017 \\ 0.102 \\ 0.396 \end{bmatrix} u_k \quad (5.2a)$$

$$y_k = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x_k \quad (5.2b)$$

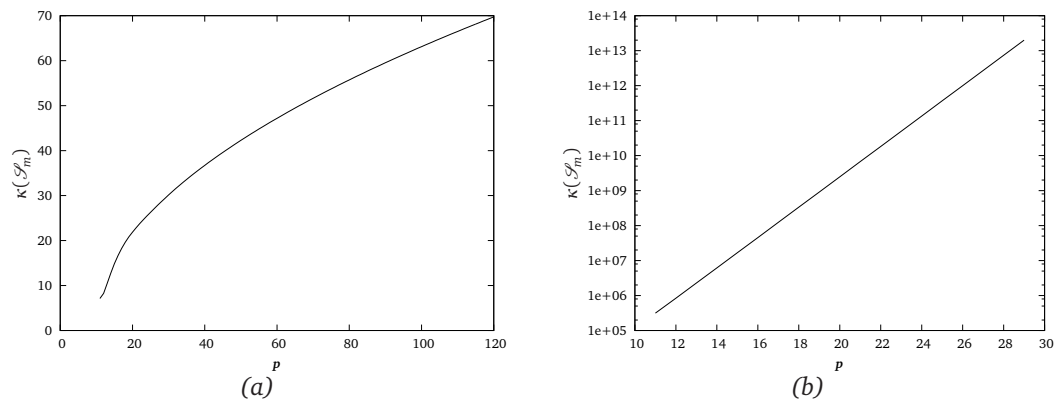
com  $x_k \in \mathbb{R}^3$  e  $u_k, y_k \in \mathbb{R}$ . Estes modelos serão referidos como LTI estável (5.1) e LTI instável (5.2). A utilização simultânea destes modelos simples permitirá comparar o desempenho das diferentes formulações quando aplicadas a sistemas estáveis e instáveis, uma vez que esta é a única característica diferenciadora entre os dois sistemas. Estes exemplos têm também a vantagem de ser possível a obtenção da solução analítica dos problemas de controlo óptimo correspondentes.

### 5.1.2 Condicionamento numérico da matriz dinâmica

Para estabelecer uma referência para os resultados obtidos neste Capítulo, foi avaliada a condição numérica da matriz dinâmica e da matriz Hessiana do problema (4.16), para os dois exemplos de demonstração considerados. Os resultados obtidos, utilizando como parâmetros de controlo  $p = 30$ ,  $m = 10$ ,  $Q_y = I_{n_y}$  e  $Q_u = I_{n_u}$ , são apresentados na Tabela 5.1.

Pode verificar-se que a matriz dinâmica do sistema instável, com os horizontes de controlo considerados, é singular, o que resulta numa matriz Hessiana também singular. Nestas condições, não é possível a solução do problema (4.16), com a precisão correspondente à representação numérica usada (precisão dupla, segundo a norma IEEE 754). Já para o sistema estável, ambas as matrizes têm números de condição baixos, sendo o da matriz Hessiana inferior devido a alguma regularização proporcionada pelo termo aditivo  $Q_2$ . Resultados semelhantes são apresentados por Santos (2001), que utiliza os mesmos exemplos para comparar o condicionamento numérico de diferentes métodos de solução do problema de controlo. O problema do mau condicionamento da matriz dinâmica agrava-se com a utilização de horizontes de previsão mais longos, o que pode ser observado na Figura 5.1, onde é apresentada a evolução do número de condição da matriz dinâmica com o tamanho do horizonte de previsão,

$p$ , para um horizonte de controlo constante,  $m = 10$ .



**Figura 5.1** Número de condição da matriz dinâmica  $\mathcal{S}_m$  em função do tamanho do horizonte de previsão, com  $m = 10$ , para os sistemas LTI (a) estável e (b) instável.

Verifica-se que, para o sistema LTI estável, o crescimento de  $\kappa(\mathcal{S}_m)$  com o tamanho do horizonte de previsão é aproximadamente linear. Para o sistema instável, este crescimento é aproximadamente exponencial para valores de  $p$  entre 10 e 29. Para valores superiores, a matriz é singular.

## 5.2 Reformulação do problema quadrático

No âmbito da formulação de controlo predictivo original, a matriz dinâmica é obtida a partir das equações do modelo expandidas para todo o horizonte de previsão, para expressar directamente a relação entre as variáveis de saída e as variáveis manipuladas, através de (4.12), eliminando neste processo as variáveis de estado do problema. A substituição desta relação no problema de controlo predictivo, permite a obtenção do problema de optimização (4.16), em que as variáveis de decisão são, exclusivamente, as variáveis manipuladas.

De forma a evitar a utilização desta matriz e os problemas associados, procuraram-se formas alternativas de representar o modelo do processo, nomeadamente utilizando as equações individuais (4.11), correspondentes à linearização deste modelo em cada um dos intervalos de amostragem. A representação conjunta destas equações para todo o horizonte de previsão resulta num sistema de equações lineares, que poderão ser consideradas em simultâneo. Uma vantagem desta representação alternativa, apesar da sua maior dimensão, é a possibilidade da aplicação de métodos robustos no seu tratamento, como os utilizados em problemas de mínimos quadrados, particularmente quando o problema é *rank deficient* ou numericamente mal condicionado (Golub et al., 1976; Golub e Van Loan, 1996).







$$b_k = \begin{bmatrix} (-2e_{k+p}^T Q_y)^T & 0 & (-2e_{k+p-1}^T Q_y)^T & 0 & \dots \\ \dots & (-2e_{k+1}^T Q_y)^T & 0 & (-2u'_{k+m-1} Q_u)^T & \dots & (-2u'_k Q_u)^T \end{bmatrix}^T;$$

$$V^{Ld} = \begin{bmatrix} (y_{k+p}^L - \bar{y}_{k+p})^T & (x_{k+p}^L - \bar{x}_{k+p})^T & \dots \\ \dots & (y_{k+1}^L - \bar{y}_{k+1})^T & (x_{k+1}^L - \bar{x}_{k+1})^T & (u_{k+m-1}^L - \bar{u}_{k+m-1})^T & \dots \\ \dots & \dots & \dots & \dots & (u_k^L - \bar{u}_k)^T \end{bmatrix}^T;$$

$$V^{Ud} = \begin{bmatrix} (y_{k+p}^U - \bar{y}_{k+p})^T & (x_{k+p}^U - \bar{x}_{k+p})^T & \dots \\ \dots & (y_{k+1}^U - \bar{y}_{k+1})^T & (x_{k+1}^U - \bar{x}_{k+1})^T & (u_{k+m-1}^U - \bar{u}_{k+m-1})^T & \dots \\ \dots & \dots & \dots & \dots & (u_k^U - \bar{u}_k)^T \end{bmatrix}^T,$$

com  $H \in \mathbb{R}^{(n_x+n_y+n_u) \times (n_x+n_y+n_u)}$  e  $b_k, V^{Ld}, V^{Ud} \in \mathbb{R}^{n_x+n_y+n_u}$ .

O número de variáveis deste problema é superior ao do problema (4.16), sendo o número de graus de liberdade idêntico. No âmbito da solução do problema de optimização (4.2), a utilização de qualquer uma das abordagens é equivalente, pelo que a utilização deste QP corresponde ainda a uma aproximação de Gauss–Newton do problema original.

## 5.2.2 Decomposição da solução

Ainda que o problema quadrático (5.4) possa ser resolvido utilizando qualquer um dos métodos aplicáveis à resolução deste tipo de problema, como os descritos na Secção 2.5.1, irá ser considerado um método baseado na decomposição da solução nos espaços característico e nulo da matriz das restrições (Gill et al., 1981; Fletcher, 1987). Esta escolha é feita de forma a que a reformulação proposta continue a possibilitar tirar partido da pequena dimensão dos problemas de optimização proporcionada pela utilização de métodos simultâneos.

Na resolução do problema, as variáveis  $V_k$  são divididas em dependentes e independentes, correspondendo cada um destes conjuntos aos espaços característicos e nulos da matriz das restrições do problema,  $M_k$ . Esta divisão corresponde a uma partição das colunas desta matriz como

$$M_k = [\mathcal{C}_k | \mathcal{N}_k]$$

em que o número de colunas de  $\mathcal{C}_k$  é igual ao número de variáveis dependentes ( $p \times (n_x + n_y)$ ) e o número de colunas de  $\mathcal{N}_k$  é igual ao número de variáveis independentes

$(m \times n_u)$ . Esta divisão das variáveis corresponde a uma decomposição da solução

$$V = W_k d_w + Z_k d_z$$

onde  $d_w \in \mathbb{R}^{p \times (n_x + n_y)}$  e  $d_z \in \mathbb{R}^{m \times n_u}$  são as componentes da solução nos espaços característico e nulo, respectivamente. A matriz  $Z_k$  deve ser tal que

$$M_k Z_k = 0$$

e as colunas de  $W_k$  devem formar uma base do espaço característico de  $M_k$ . Considerando a partição das colunas de  $M_k$  apresentada acima, estas matrizes podem ser (Fletcher, 1987):

$$Z_k = \begin{bmatrix} -\mathcal{C}_k^{-1} \mathcal{N}_k \\ I \end{bmatrix}, \quad W_k = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad (5.5)$$

A componente da solução no espaço característico, correspondente às variáveis dependentes, é completamente determinada pelas restrições, sendo  $W_k d_w$  a solução básica de  $M_k V = a_k$ . O passo no espaço nulo é a solução do QP (Cervantes e Biegler, 2000)

$$\min_{d_z} (Z_k^T b_k + 2Z_k^T H W_k d_w)^T d_z + d_z^T (Z_k^T H Z_k) d_z \quad (5.6a)$$

$$\text{s.a. } V^{Ld} - V_k - W_k d_w \leq Z_k d_z \leq V^{Ud} - V_k - W_k d_w \quad (5.6b)$$

onde  $Z_k^T H Z_k$  é a Hessiana projectada no espaço nulo das restrições. A dimensão deste problema é igual à do problema (4.16).

Um aspecto importante da aplicação deste método à solução dos problemas quadráticos gerados em cada iteração de um método sequencial é o facto de a linearização do modelo ser feita em torno de uma trajectória viável. Logo, em qualquer iteração,  $\Delta x_k$  é sempre nulo e também o lado direito das restrições de igualdade,  $a_k$ , é sempre nulo, resultando daqui que a componente do passo no espaço característico,  $d_w$ , é sempre nula. Em consequência, não só não é necessário o seu cálculo como também não é necessário obter, ou aproximar, o termo cruzado no problema (5.9), uma vez que este é nulo. Contudo, por uma questão de generalidade, na discussão que se segue este termo cruzado será mantido.

Relativamente à escolha das variáveis independentes, importa lembrar que, em algumas situações, a existência de restrições activas (nomeadamente variáveis de entrada ou de saída saturadas) pode causar instabilidade em ciclo fechado, devido à necessidade de utilização de determinados conjuntos de variáveis de decisão, que se podem encontrar limitados pelas restrições do problema. Este problema ocorrerá qualquer que seja o método utilizado para a solução do problema de MPC. A sua ocorrência depende apenas da natureza do sistema e das suas restrições. Este problema apenas é ultrapassado pela utilização de *soft constraints* (Oliveira, 1994), não sendo considerado neste trabalho.

### Partição das variáveis

Uma partição adequada das variáveis do problema em dependentes e independentes é crucial para a estabilidade numérica da solução do problema. Esta partição deve ser efectuada de forma a que as matrizes  $\mathcal{C}_k$  e  $Z_k$  e a Hessiana projectada,  $Z_k^T H Z_k$ , sejam não-singulares e, tanto quanto possível, bem condicionadas. Na discussão que se segue, para simplificar a notação, serão utilizadas as quantidades

$$n_e = p \times (n_x + n_y) \quad \text{e} \quad n_v = n_e + m \times n_u$$

correspondentes aos número de restrições e ao número de variáveis, respectivamente. Será também utilizada a definição de característica da matriz (Golub e Van Loan, 1996)

$$r = \text{car}(M_k)$$

igual ao número máximo de linhas ou colunas de  $M_k$  linearmente independentes. De acordo com esta definição, a matriz  $M_k$  é não-singular quando  $r = n_e$  e singular quando  $r < n_e$ .

Quando a matriz  $M_k$  é não singular, pode ser considerado um número de variáveis dependentes igual ao número de restrições, sendo neste caso o número de variáveis independentes  $(n_v - n_e) = m \times n_u$ . A componente no espaço característico de  $M_k$  da solução de (5.4) é

$$d_w = [M_k W_k]^{-1} a_k = \mathcal{C}_k^{-1} a_k$$

ou seja,  $d_w$  é a solução do sistema linear:

$$\mathcal{C}_k d_w = a_k. \quad (5.7)$$

Facilmente verificamos que quando  $W_k$  é escolhido como em (5.5),  $W_k d_w$  é a solução básica de  $M_k V = a_k$ .

No problema de controlo, a escolha natural para as variáveis independentes serão as variáveis manipuladas  $\Delta U$ , contendo a matriz  $\mathcal{C}_k$  as primeiras  $n_e$  colunas de  $M_k$ , correspondentes às variáveis medidas e de estado. Esta matriz é triangular superior, com todos os elementos da diagonal não-nulos e, por isso, estruturalmente não-singular. A utilização desta partição corresponde à solução do sistema de equações (5.3) de forma a explicitar a relação entre as variáveis medidas e de estados e as variáveis manipuláveis, o que, com facilidade, é possível fazer simbolicamente dada a sua estrutura. Se para além disso, forem também eliminadas as variáveis de estado, obtém-se uma relação semelhante a (4.12). Nestas condições, a utilização da decomposição da solução é equivalente à utilização da matriz dinâmica, resultando nos mesmos problemas de mau condicionamento numérico.

A escolha da partição de variáveis em dependentes e independentes possibilita que

sejam efectuadas trocas entre os dois conjuntos de variáveis, o que será reflectido nas colunas de  $M_k$  que farão parte de  $\mathcal{C}_k$ . De forma a melhorar a robustez da solução do problema quadrático, estas colunas deverão ser escolhidas de forma a garantir o bom condicionamento de  $\mathcal{C}_k$ , pelo que o problema da selecção do conjunto de variáveis independentes mais adequado para a solução do problema de optimização passa por escolher qual o conjunto de colunas que permite um melhor condicionamento desta matriz. Esta escolha pode ser feita recorrendo à factorização QR com pivotagem, um método cuja robustez numérica é bem conhecida na solução de problemas de mínimos quadrados (Golub e Van Loan, 1996). Aplicando esta pivotagem à matriz  $M_k$ , quando esta é não-singular, obtém-se

$$M_k \Pi = QR = Q[R_1 | R_2]$$

onde  $Q \in \mathbb{R}^{n_e \times n_e}$  é uma matriz ortogonal,  $R_1 \in \mathbb{R}^{n_e \times n_e}$  é uma matriz triangular superior,  $R_2 \in \mathbb{R}^{n_e \times (n_v - n_e)}$  é uma matriz densa e  $\Pi \in \mathbb{N}^{n_v \times n_v}$  é uma matriz de permutações. A matriz  $R_1$  é sempre não-singular e a pivotagem das colunas é feita de forma a que seja bem condicionada. A matriz de permutações obtida na factorização QR pode ser utilizada para permutar as variáveis de decisão do QP (5.4):

$$\hat{V} = \Pi V$$

Utilizando esta permutação das variáveis, pode ser escrito um QP equivalente a (5.4), com a forma:

$$\min_{\hat{V}} \quad g_k^T \hat{V} + \hat{V}^T H \hat{V} \quad (5.8a)$$

$$\text{s.a} \quad QR \hat{V} = a_k \quad (5.8b)$$

$$\hat{V}^{Ld} \leq \hat{V} \leq \hat{V}^{Ud} \quad (5.8c)$$

onde

$$g_k^T = b_k^T \Pi$$

$$\hat{H} = \Pi^T H \Pi$$

$$\hat{V}^{Ld} = \Pi^T V^{Ld}$$

$$\hat{V}^{Ud} = \Pi^T V^{Ud}$$

Neste problema, a matriz das restrições é QR, pelo que as matrizes  $\mathcal{C}_k$  e  $\mathcal{N}_k$  serão

$$\mathcal{C}_k = QR_1 \quad \mathcal{N}_k = QR_2$$

ou seja, matriz  $\mathcal{C}_k$  será não-singular e bem condicionada. As matrizes  $Z_k$  e  $W_k$  podem

ser obtidas por

$$Z_k = \begin{bmatrix} -R_1^{-1}R_2 \\ I \end{bmatrix}, \quad W_k = \begin{bmatrix} I_{n_e} \\ 0 \end{bmatrix}$$

A componente da solução no espaço característico das restrições é

$$d_w = R_1^{-1}Q^T a_k$$

e a componente no espaço nulo é a solução do QP:

$$\min_{d_z \in \mathbb{R}^{(n_v - n_e)}} (Z_k^T g_k + 2Z_k^T \hat{H} W_k d_w)^T d_z + d_z^T (Z_k^T \hat{H} Z_k) d_z \quad (5.9a)$$

$$\text{s.a } \hat{V}_{ld} - \hat{V}_k - W_k d_w \leq Z_k d_z \leq \hat{V}_{ud} - \hat{V}_k - W_k d_w \quad (5.9b)$$

Na ausência de restrições activas,  $d_z$  pode ser obtido pela solução do sistema linear simétrico:

$$(Z_k^T \hat{H} Z_k) d_z = -\frac{1}{2} (Z_k^T g_k + 2Z_k^T \hat{H} W_k d_w) \quad (5.10)$$

Como já referido anteriormente, na aplicação desta decomposição da solução em métodos sequenciais, o passo no espaço característico,  $d_w$ , será nulo, em consequência de apenas serem consideradas trajectórias viáveis. Daqui resulta uma simplificação considerável deste problema, nomeadamente a anulação do termo cruzado  $2Z_k^T \hat{H} W_k d_w$ .

**Matriz das restrições singular** Quando  $\text{car}(M_k) < n_e$ , a matriz das restrições é numericamente singular, sendo necessárias algumas alterações no procedimento de partição das variáveis descrito anteriormente. Nesta situação, o resultado da factorização QR com pivotagem de  $M_k$  é (Golub e Van Loan, 1996)

$$M_k \Pi = QR = \begin{bmatrix} Q1 & Q2 \\ r & (n_e - r) \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix} \begin{matrix} r \\ (n_e - r) \end{matrix} \quad (5.11)$$

onde  $Q \in \mathbb{R}^{n_e \times n_e}$  é uma matriz ortogonal,  $R_{11} \in \mathbb{R}^{r \times r}$  é uma matriz triangular superior não-singular e  $\Pi$  é uma matriz de permutações. À semelhança da situação em que  $M_k$  é não-singular, a matriz das permutações pode continuar a ser utilizada para obter um QP na forma (5.8), dividindo a solução deste problema nos componentes nulo e característico.

A componente  $W_k d_w$ , correspondente à solução básica das restrições, tem no máximo  $r$  elementos não-nulo e pode ser obtida por (Golub e Van Loan, 1996):

$$W_k d_w = \begin{bmatrix} R_{11}^{-1} c \\ 0 \end{bmatrix} \begin{matrix} r \\ (n_v - r) \end{matrix}$$

onde  $c$  é obtido a partir de

$$Q^T a_k = \begin{bmatrix} c \\ d \end{bmatrix} \begin{matrix} r \\ (n_e - r) \end{matrix}$$

ou seja,  $c$  contém os primeiros  $r$  elementos de  $Q^T a_k$ . Nesta situação, a matriz  $W_k$  pode ser definida como

$$W_k = \begin{bmatrix} I_r \\ 0 \end{bmatrix}$$

e:

$$d_w = R_{11}^{-1} c$$

A inversa de  $R_{11}$  existe, uma vez que esta matriz é não-singular. Verifica-se que  $d_w \in \mathbb{R}^r$ , ou seja, quando a matriz das restrições é singular com  $\text{car}(M_k) = r < n_e$  apenas podem existir  $r$  variáveis dependentes, logo existirão  $n_v - r$  variáveis independentes e  $d_z \in \mathbb{R}^{n_v - r}$ . Para obter a matriz  $Z_k$  é útil verificar, a partir de (5.11), que

$$QR = \begin{bmatrix} Q_1 R_{11} & Q_1 R_{12} \end{bmatrix}$$

e também que

$$R_{11}^{-1} Q_1^T Q_1 R_{11} = R_{11}^{-1} I_r R_{11} = I_r$$

uma vez que  $Q_1$  é um subconjunto das colunas da matriz ortogonal  $Q$ . Então, a matriz  $Z_k$  pode ser construída como (Bischof e Quintana-Ortí, 1998):

$$Z_k = \begin{bmatrix} -R_{11}^{-1} R_{12} \\ I \end{bmatrix} \begin{matrix} r \\ (n_v - r) \end{matrix},$$

uma vez que

$$\begin{aligned} \begin{bmatrix} Q_1 R_{11} & Q_1 R_{12} \end{bmatrix} \begin{bmatrix} -R_{11}^{-1} R_{12} \\ I \end{bmatrix} &= \begin{bmatrix} -Q_1 R_{11} R_{11}^{-1} R_{12} + Q_1 R_{12} \end{bmatrix} \\ &= 0 \in \mathbb{R}^{n_e \times (n_v - r)} \end{aligned}$$

Depois de conhecidas as matrizes  $Z_k$  e  $W_k$  e a componente da solução no espaço característico,  $d_w$ , pode obter-se a componente no espaço nulo,  $d_z$ , resolvendo (5.9), agora com  $d_z \in \mathbb{R}^{n_v - r}$ .

É ainda importante referir que a decomposição (5.11) pode não ser totalmente eficaz na detecção de matrizes singulares, ou numericamente singulares (quase singulares). Nesta situação pode ser preferível a utilização do método proposto por Golub et al. (1976) para a escolha do subconjunto de variáveis a utilizar na solução de problemas de mínimos quadrados. O método proposto consiste na utilização da decomposição SVD para a detecção da singularidade e para a determinação da característica da matriz. A decomposição QR é depois aplicada aos factores obtidos para selecção

do subconjunto de variáveis, de forma a originar um problema numericamente bem condicionado.

### 5.3 Adaptação do horizonte de previsão

A reformulação do problema de controlo apresentada na Secção anterior, permite resolver um dos problemas frequentes nos métodos sequenciais, o mau condicionamento numérico do problema de optimização. A sua utilização depende, no entanto, da disponibilidade da informação das derivadas de primeira ordem em todo o horizonte de previsão, que no caso genérico de DAE's, podem ser obtidas por integração simultânea das equações do modelo e das sensibilidades. Em sistemas instáveis a obtenção deste informação pode ser problemática, uma vez que esta integração é feita em ciclo aberto. Ou seja, é resolvido como um problema às condições iniciais sem que, durante a integração, haja qualquer realimentação por parte do algoritmo de controlo. Para sistemas instáveis, esta integração poderá não ser possível para uma dada trajetória de controlo, por serem atingidas condições em que a utilização do modelo não é viável, o que será tão mais provável quanto maior o horizonte de previsão utilizado.

Outra questão importante é o facto de, durante a integração das equações do modelo e das sensibilidades, o erro apenas ser controlado localmente (Kiehl, 1999). Em particular no caso de sistemas instáveis, estes erros poderão acumular-se, sendo o resultado da integração imprevisível. Mesmo que a integração das equações do modelo e das sensibilidades seja possível, a sua precisão poderá não ser suficiente para uma solução adequada do problema de optimização. Novamente, este problema será tão mais grave quanto maior for o horizonte de previsão.

Para que, nas situações atrás referidas, continue a ser possível a utilização da abordagem sequencial, esta poderá ser adaptada de forma semelhante à proposta por Michalik et al. (2009) para problemas de estimação de parâmetros, com uma estratégia designada *incremental single shooting* (ISS). Esta estratégia surge da observação que as falhas que ocorrem na integração das equações do modelo devidas à utilização de estimativas erradas dos parâmetros, ocorrem geralmente quando a integração já está avançada e não nos instantes iniciais do horizonte de estimação. A ideia base da metodologia ISS é evitar a integração das equações do modelo em todo horizonte de estimação enquanto as estimativas dos parâmetros forem incertas, resolvendo o problema em horizontes mais pequenos que vão sendo sucessivamente aumentados. Segundo os autores, a forma como o tamanho dos intervalos é escolhida é importante do ponto de vista da robustez do método. Uma opção possível será primeiro tentar integrar as equações do modelo em todo o horizonte. Se esta integração for possível, o método pode prosseguir sem que seja necessária a utilização de horizontes mais pequenos. Se ocorrer uma falha no instante  $t_{\text{erro}}$ , então o tamanho do primeiro hori-



zonte em que o problema é resolvido pode ser escolhido como  $\alpha t_{\text{erro}}$ , com um valor de  $\alpha$  entre 0 e 1. A estratégia utilizada pelos autores é considerar intervalos tais que, na iteração  $i$ , sejam utilizados  $2^{i-1}$  pontos, resultando na utilização de intervalos que, em cada iteração, contêm o dobro dos valores utilizados na iteração anterior.

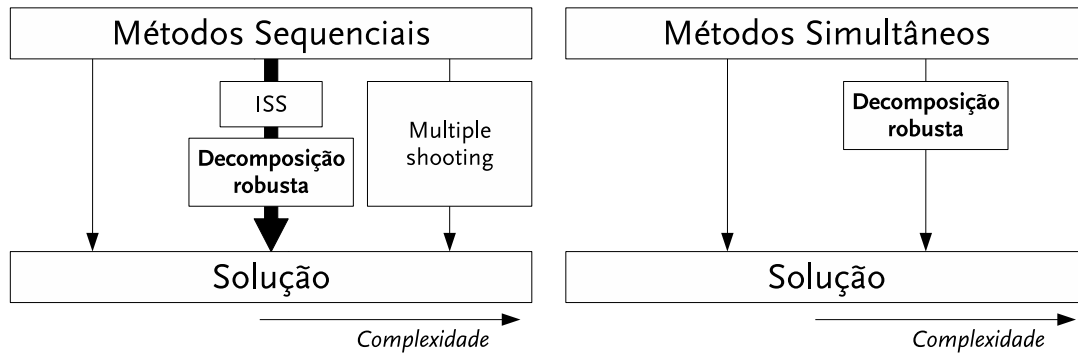
Os resultados apresentados pelos autores demonstram uma maior robustez do método ISS, relativamente ao método de *single shooting*, e também melhor capacidade de convergir para o óptimo global. A utilização deste método pode resultar num maior número de iterações, sendo que uma parte destas iterações corresponde à solução do problema em horizontes pequenos e, portanto, a problemas de pequena dimensão. Outro aspecto referido é a facilidade com que este método pode ser implementado, utilizando qualquer implementação de um método sequencial.

A aplicação de uma estratégia semelhante a problemas de controlo predictivo surge de forma natural. Neste tipo de problemas, as falhas na integração das equações do modelo surgem em consequência da utilização de trajectórias de controlo que não são capazes de estabilizar o sistema em todo o horizonte de previsão. Se estas falhas ocorrerem, o problema pode primeiro ser resolvido utilizando um horizonte de previsão mais curto, sendo que, de forma geral, a solução de controlo obtida permite estabilizar o sistema para além do fim do horizonte de previsão utilizado. Tirando partido deste facto, a solução obtida neste horizonte pode depois ser utilizada como solução de um problema de controlo com um horizonte mais longo.

O procedimento proposto para lidar com falhas na integração das equações do modelo em problemas de controlo predictivo é:

1. num instante  $t_0$ , dada uma trajectória de controlo inicial, resolver o problema de controlo no horizonte  $]t_0, t_0 + p_{\text{ref}}T]$ ; se a integração falhar passar para 2;
2. se ocorreu uma falha no instante  $t = t_{\text{erro}}$ , reduzir o horizonte de previsão de forma a que  $p_1 T < \alpha t_{\text{erro}}$ , com  $\alpha \in ]0, 1]$  e  $p_1 > m$ ;
3. resolver o problema de controlo no horizonte de previsão  $p_1$ ;
4. enquanto  $p_i < p_{\text{ref}}$ ,  $i > 1$ :
  - (a) considerar um horizonte de previsão  $p_i > p_{i-1}$ ;
  - (b) resolver o problema de controlo no horizonte  $p_i$ , utilizando como trajectória inicial a solução de controlo obtida com o horizonte  $p_{i-1}$ ;

O número de horizontes,  $p_i$  a considerar neste procedimento deverá ser um compromisso entre a robustez e o esforço de cálculo necessário para obter a solução. Para a implementação deste tipo de estratégia, a utilização dos algoritmos de controlo predictivo considerados neste trabalho tem a vantagem de a dimensão do problema ser independente do tamanho do horizonte de previsão, enquanto o tamanho dos horizontes de previsão utilizados for superior ao tamanho do horizonte de controlo



**Figura 5.3** Estratégias de solução de problemas de otimização dinâmica.

( $p_{\text{ref}} > p_i > m$ ). Nestes casos, é mais simples a aplicação de uma estratégia como a aqui descrita.

## 5.4 Comparação com outras estratégias

Para comparar a eficácia da abordagem proposta nas Secções anteriores, é importante enquadrá-la quer no âmbito da aplicação dos métodos sequenciais, quer no âmbito dos métodos directos de solução de problemas de otimização dinâmica em geral. Na Figura 5.3 é representada, de forma esquemática, uma possível integração desta com alguns dos principais métodos de otimização dinâmica discutidos anteriormente. A abordagem proposta neste trabalho é a correspondente à utilização de uma decomposição robusta e de horizontes de previsão adaptáveis (ISS).

Relativamente aos métodos sequenciais “puros”, como a formulação de controlo predictivo apresentada na Secção 4.1, a principal diferença da abordagem proposta reside na utilização de uma decomposição robusta da matriz de restrições resultante da linearização das equações do modelo no horizonte de previsão. A robustez desta decomposição resulta da utilização de pivotagem das colunas da matriz  $M_k$  (Figura 5.2). Esta pivotagem determina a forma como as variáveis do problema são particionadas. Quando não é feita pivotagem, as variáveis independentes do problema de otimização serão sempre as variáveis manipuláveis, obtidas pela solução do problema (5.6). As restantes variáveis são obtidas por  $\mathcal{C}_k^{-1} \mathcal{N}_k d_z$ .

Por outro lado, quando é utilizada pivotagem, as variáveis independentes do problema podem ser qualquer combinação de variáveis manipuláveis, medidas ou de estado, em qualquer instante de amostragem do horizonte de previsão. Neste caso, o algoritmo tem liberdade para escolher o conjunto de variáveis independentes que resultem num melhor condicionamento numérico do problema. Quando é utilizada a factorização QR com pivotagem, são primeiro escolhidos os valores das variáveis independentes, pela solução do problema (5.8), que corresponde ao passo no espaço

nulo da matriz das restrições,  $d_z$ . A partir destas, são depois obtidas as variáveis dependentes através de  $-R_1^{-1}R_2d_z$ .

A utilização de pivotagem permite resolver o problema do mau condicionamento numérico do problema, mas continua a depender da integração das equações do modelo, em ciclo aberto, até ao final do horizonte de previsão. Esta integração pode nem sempre ser possível, particularmente no casos de sistemas instáveis e horizontes de grande dimensão, em que também podem ocorrer problemas com a precisão das soluções obtidas. A utilização da estratégia de adaptação do horizonte de previsão pode ultrapassar estes problemas ao considerar, quando necessário, a resolução do problema de controlo em horizontes sucessivamente crescentes. A utilização sucessiva da solução de controlo obtida num horizonte mais pequeno permite uma estabilização que poderá ser suficiente para a integração das equações do modelo em horizontes progressivamente maiores, até ser atingida a dimensão pretendida do horizonte de previsão. Para esta estabilização ser mais eficiente poderão ser ainda consideradas condições adicionais, tais como restrições terminais nos vários horizontes considerados.

Nas estratégias do tipo *multiple shooting*, quer o problema do condicionamento numérico quer o problema da possível falha da integração do modelo em ciclo aberto, são ultrapassados pela integração das equações do modelo, de forma independente, em cada um dos intervalos de amostragem do horizonte de previsão. Desta forma, as equações do modelo apenas são integradas em horizontes pequenos, diminuindo o mau condicionamento numérico do modelo linearizado que, como já referido, aumenta exponencialmente com o tamanho do horizonte. A utilização de horizontes curtos, quando utilizadas condições iniciais adequadas, também permite reduzir significativamente a probabilidade de ocorrência de falhas na integração das equações do modelo. A implementação desta estratégia resulta em problemas de maior dimensão. Esta maior complexidade deve-se à necessidade de considerar variáveis adicionais que, como descrito na Secção 2.3.2, representam as condições iniciais para a integração do modelo em cada um dos intervalos de amostragem. É também necessário considerar restrições adicionais de continuidade em todos os instantes de amostragem no interior do horizonte de previsão, para garantir que na solução do problema de controlo, a trajectória das variáveis de estado obtida pela solução das equações do modelo seja contínua. Esta complexidade será tanto maior quanto menores forem os intervalos de *shooting* considerados.

Nos métodos simultâneos não ocorrem as dificuldades associadas à integração das equações do modelo em ciclo aberto, uma vez que esta não é feita, sendo os problemas de controlo e de solução das equações do modelo resolvidos simultaneamente (uma única vez). Deve, no entanto, ser também considerada a possibilidade do surgimento de problemas mal condicionados nesta classe de métodos, nomeadamente quando o problema de optimização é resolvido por métodos que decompõem a solução. Quando são utilizados métodos de colocação, como os descritos na Secção 2.3.2,

as variáveis de decisão do problema resolvido nos métodos simultâneos são os coeficientes (parâmetros) das funções base utilizadas na parametrização da solução. No caso da utilização de polinómios interpoladores de Lagrange, estes coeficientes correspondem a todas as variáveis manipuláveis, medidas e de estado em todos os pontos de colocação. Daqui resulta que a representação das equações do modelo linearizado será uma matriz diagonal por blocos, com estrutura semelhante à da matriz  $M_k$ . Sendo as representações do modelo semelhantes, é de esperar que em alguns casos o problema de optimização resultante possa também apresentar mau condicionamento numérico (Biegler, 2010), sobretudo quando o problema é resolvido de forma a tirar partido da estrutura e esparsidade do problema. Assim, em alguns casos, a decomposição robusta pode também ser útil nesta classe de métodos, sendo no entanto necessária a factorização com pivotagem de matrizes com muito maior dimensão do que as correspondentes à utilização de métodos sequenciais.

A escolha da abordagem mais adequada para o controlo predictivo de um dado sistema depende de vários factores, entre os quais a forma como o modelo está (ou será) implementado e a sua estabilidade dinâmica. Enquanto que para a generalidade dos sistemas estáveis os métodos sequenciais podem resolver o problema directamente, no caso de sistemas instáveis ou *stiff* será frequentemente necessário o recurso à utilização de estratégias mais complexas, como a abordagem descrita neste Capítulo. No caso dos métodos simultâneos, esta necessidade é mitigada devido à forma simultânea como as equações do modelo são resolvidas, podendo, no entanto, ainda ser conveniente a utilização de decomposições robustas, de forma a melhorar a robustez numérica deste tipo de formulação.

Finalmente, e tal como referido na Secção 5.2.2, é possível ainda constatar que quando os problemas de estabilidade são causados por restrições activas, que nesta situação condicionam severamente quais as variáveis que podem ser escolhidas como independentes, deverá ser procurado outro tipo de soluções para ultrapassar estes problemas, como referido na Secção 5.2.

## 5.5 Resultados de aplicação aos exemplos de teste

Nesta secção são apresentados os resultados obtidos pela aplicação da reformulação proposta aos dois sistemas LTI descritos na Secção 5.1.1. Primeiro é analisado o condicionamento numérico do problema, de forma a verificar as melhorias introduzidas pelas permutações de variáveis. De seguida são identificadas as permutações feitas durante a factorização da matriz das restrições, sendo discutidos os seus efeitos sobre a partição das variáveis do problema ao nível do problema quadrático. Finalmente são apresentados alguns resultados de simulação.

**Tabela 5.2** Condição numérica das diferentes matrizes envolvidas na obtenção da solução do problema de controlo do sistema LTI estável (5.1).

	$\kappa(R_1)$	$\kappa(Z)$	$\kappa(H_z)$
QR	23.0	4.26	8.39
QR c/ pivotagem	43.4	6.29	18.3

**Tabela 5.3** Condição numérica das diferentes matrizes envolvidas na obtenção da solução do problema de controlo do sistema LTI instável (5.2).

	$\kappa(R_1)$	$\kappa(Z)$	$\kappa(H_z)$
QR	$3.59 \times 10^{13}$	$1.19 \times 10^{12}$	$4.89 \times 10^{22}$
QR c/ pivotagem	34.6	1.42	21.9

### 5.5.1 Condicionamento numérico

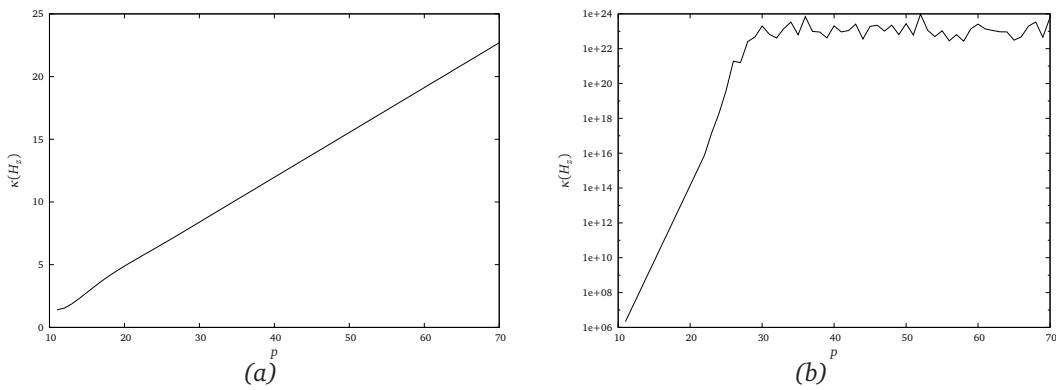
Na Tabela 5.2 são apresentados os números de condição das matrizes  $R_1$ ,  $Z$  e Hessiana projectada,  $H_z$ , para o sistema LTI estável. Os parâmetros de controlo utilizados na obtenção destes resultados foram  $p = 30$ ,  $m = 10$ ,  $Q_y = I_{n_y}$  e  $Q_u = I_{n_u}$ . As factorizações QR foram obtidas utilizando as rotinas DGEQRF, sem pivotagem, e DGEQP3, com pivotagem, ambas parte da biblioteca LAPACK (Anderson et al., 1999).

Pode verificar-se que nenhuma das matrizes é mal-condicionada, o que seria previsível dada a estabilidade em ciclo aberto do sistema. A utilização de pivotagem resulta num aumento ligeiro do número de condição. É também de notar a coincidência entre o valor do número de condição da matriz Hessiana projectada sem pivotagem e o número de condição da matriz dinâmica obtida com os mesmos parâmetros de controlo, apresentado na Tabela 5.1. Esta coincidência resulta da equivalência entre as duas abordagens.

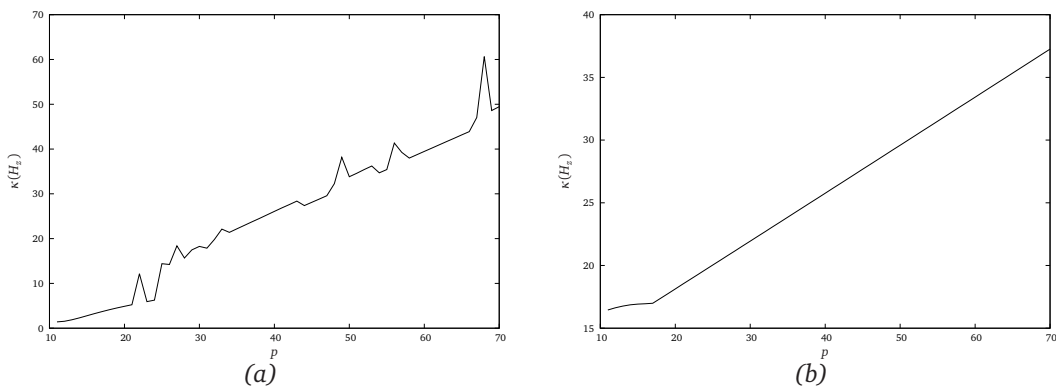
Na Tabela 5.3 são apresentados os números de condição das matrizes  $R_1$ ,  $Z$  e da Hessiana projectada,  $H_z$ , para o sistema LTI instável. Os parâmetros de controlo utilizados são os mesmos utilizados com o sistema estável.

Pode verificar-se que, quando não é feita pivotagem, todas as matrizes são severamente mal-condicionadas, particularmente a Hessiana projectada. Este mau condicionamento agravar-se-à com a utilização de horizontes de previsão mais longos. A utilização da pivotagem permite reduzir drasticamente o número de condição desta matriz, sendo conseguidos valores com a mesma magnitude do sistema LTI estável.

De forma a melhor avaliar a robustez conseguida com a utilização de pivotagem, o número de condição da Hessiana reduzida foi calculado para diferentes tamanhos do horizonte de previsão,  $p$ . Na Figura 5.4 é apresentada esta variação para os sistemas LTI estável e instável, quando não é utilizada pivotagem.



**Figura 5.4** Número de condição da matriz  $H_z$  do problema sem pivotagem, em função do tamanho do horizonte de previsão, com  $m = 10$ , para os sistemas LTI (a) estável e (b) instável.



**Figura 5.5** Número de condição da matriz  $H_z$  do problema com pivotagem, em função do tamanho do horizonte de previsão, com  $m = 10$ , para os sistemas LTI (a) estável e (b) instável.

Para o sistema estável, o crescimento do número de condição da Hessiana projectada com o tamanho do horizonte de previsão é aproximadamente linear, sendo a matriz sempre bem condicionada para qualquer um dos tamanhos do horizonte considerados. Já para o sistema instável, este crescimento é exponencial até próximo de  $p = 30$ , o que está de acordo com as estimativas apresentadas por Santos (2001). A partir daqui os valores do número de condição não apresentam nenhuma tendência definida, o que se deverá a erros numéricos no seu cálculo.

Na Figura 5.4 é apresentada esta variação para os sistemas LTI estável e instável, quando é utilizada pivotagem na partição das variáveis do problema.

O crescimento observado para o sistema LTI estável continua a ser linear, ainda que ligeiramente mais rápido do que quando não é utilizada pivotagem. No caso do sistema instável, a utilização da pivotagem consegue que o crescimento do número de condição de  $H_z$  deixe de ser exponencial e passe a ser aproximadamente linear, sendo a matriz bem condicionada para qualquer um dos tamanhos do horizonte de previsão

considerados.

### 5.5.2 Permutações de variáveis

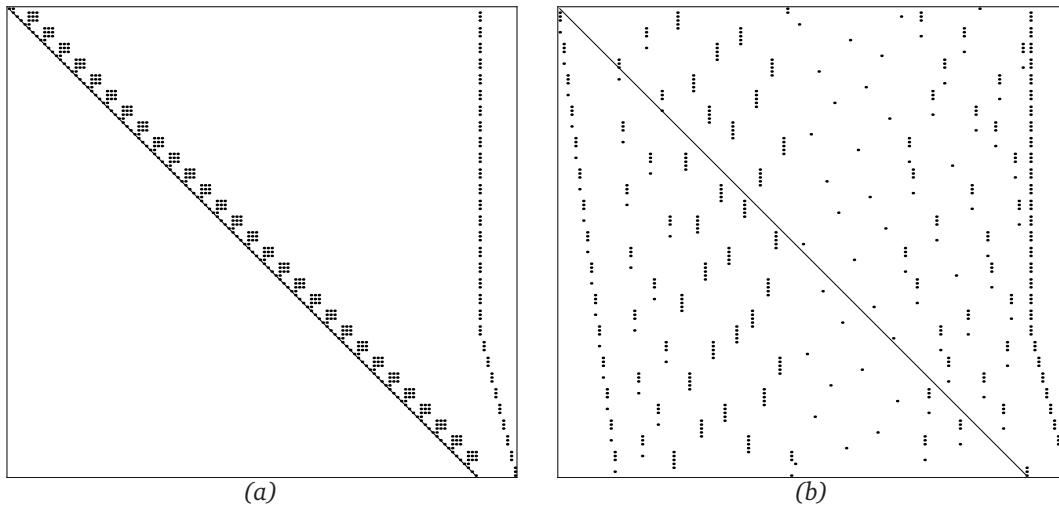
Como já referido, a utilização da factorização com pivotagem determina a selecção dos conjuntos de variáveis dependentes e independentes utilizados na solução do problema quadrático. Dada a sua dimensão, no caso dos sistemas LTI torna-se mais simples identificar qual o conjunto de variáveis independentes utilizado. Por exemplo, no caso do sistema LTI estável não é feita qualquer permutação o que corresponde à utilização das variáveis manipuladas, em todo os instantes de amostragem do horizonte de controlo, como variáveis independentes do problema de optimização.

No caso do sistema LTI instável, são efectuadas permutações, de forma a que a variável de estado correspondente ao modo instável no último instante de amostragem do horizonte de previsão deixe de ser uma variável dependente e passe a ser uma variável independente, passando a variável manipulada no primeiro instante de amostragem a ser uma variável dependente. Esta alteração foi observada para qualquer tamanho dos horizontes de previsão e de manipulação. Na representação (5.2) o modo instável é a variável de estado  $x_3$ . Quando esta variável é colocada em posições diferentes no vector das variáveis de estado, observa-se que as permutações efectuadas são sempre de forma a que esta, no último instante de amostragem, seja uma variável independente. Este facto demonstra que a aplicação da factorização QR com pivotagem é sempre capaz de isolar este comportamento.

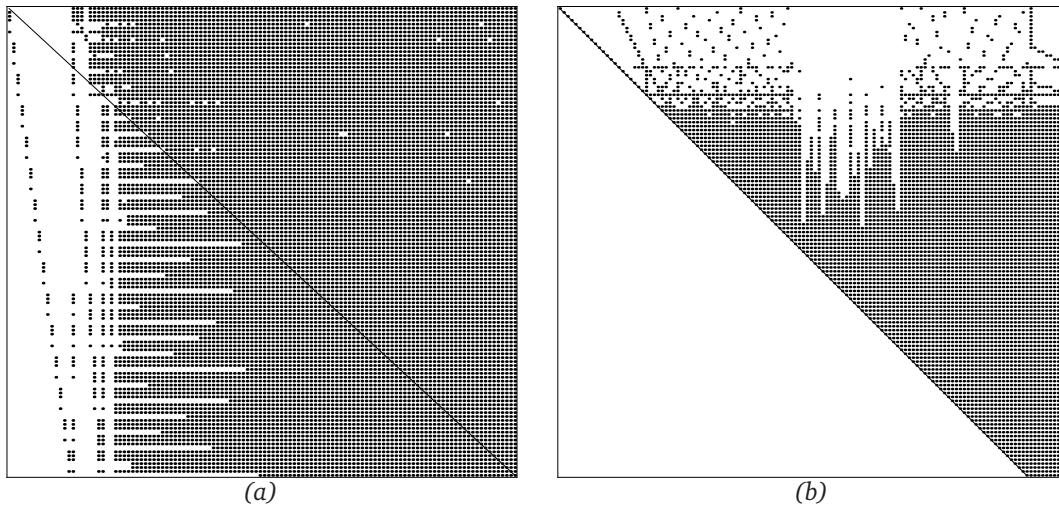
É também possível estabelecer uma analogia entre a permuta de variáveis observada e a utilização de restrições terminais (*end point constraints*). Estas restrições fixam as condições que devem ser atingidas no final do horizonte de previsão e permitem melhorar a robustez dos métodos de controlo predictivo com horizontes regressivos (Mayne e Michalska, 1990; Michalska e Mayne, 1993), sendo geralmente expressas em função das variáveis de estado como, por exemplo,  $x(t_0 + pT) = 0$ . Quando a variável de estado no último instante de amostragem, que coincide com o final do horizonte de previsão, é utilizada como variável independente no problema de optimização, o seu valor é escolhido pelo algoritmo, sendo depois as variáveis dependentes (entre as quais uma variável manipulada) calculadas de forma a atingir o valor estabelecido. Facilmente se conclui que, na ausência de restrições activas e de alterações nos *set-points*, o valor desta variável que corresponde ao valor óptimo da função objectivo em (5.4) é zero, pelo que, sempre que possível, será este o valor escolhido pelo algoritmo de optimização. Fixando, desta maneira, o valor final da variável de estado correspondente ao modo instável do sistema, evita-se a sua propagação com o tempo, estabilizando o sistema.

Deve ainda ser referido que as operações de pivotagem realizadas, desejáveis do ponto de vista da garantia da estabilidade numérica, têm efeitos indesejáveis na estrutura





**Figura 5.6** Estrutura da matriz  $M_k$  do sistema LTI instável, com horizontes  $p = 30$  e  $m = 5$ ; (a) sem pivotagem; (b) com pivotagem.



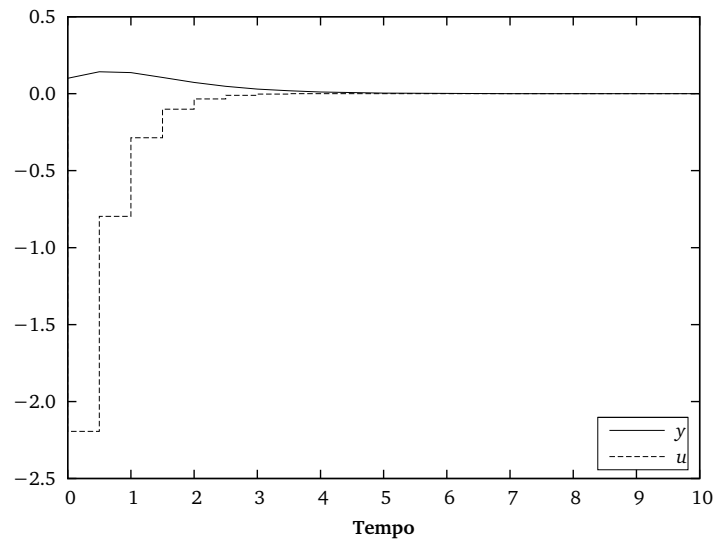
**Figura 5.7** Estrutura das matrizes (a)  $Q$  e (b)  $R$  resultantes da factorização com pivotagem da matriz  $M_k$  do sistema LTI instável, com  $p = 30$  e  $m = 10$ .

e esparsidade da matriz das restrições,  $M_k$ , e das suas factorizações. Estes efeitos são ilustrados na Figura 5.6, onde é representada a estrutura de  $M_k$ , com  $p = 30$  e  $m = 10$ , sem pivotagem e depois de aplicada a pivotagem obtida no caso do sistema LTI instável.

A factorização  $QR$  da matriz pivotada também resulta, geralmente, em matrizes densas, como pode ser verificado na Figura 5.7, onde são apresentadas as estruturas das matrizes  $Q$  e  $R$  obtidas por factorização da matriz  $M_k$  correspondente ao sistema LTI instável, com  $p = 30$  e  $m = 10$ .

Verifica-se um aumento significativo do número de elementos não nulos devido à utilização de pivotagem. Em comparação, no caso em que não é feita qualquer pivotagem,





**Figura 5.8** Perfil de controle obtido utilizando a factorização  $QR$  com pivotagem para o sistema LTI instável.

a matriz  $Q$  é simplesmente uma matriz identidade, de dimensão igual ao número de equações, e a matriz  $R$  é igual a  $M_k$ .

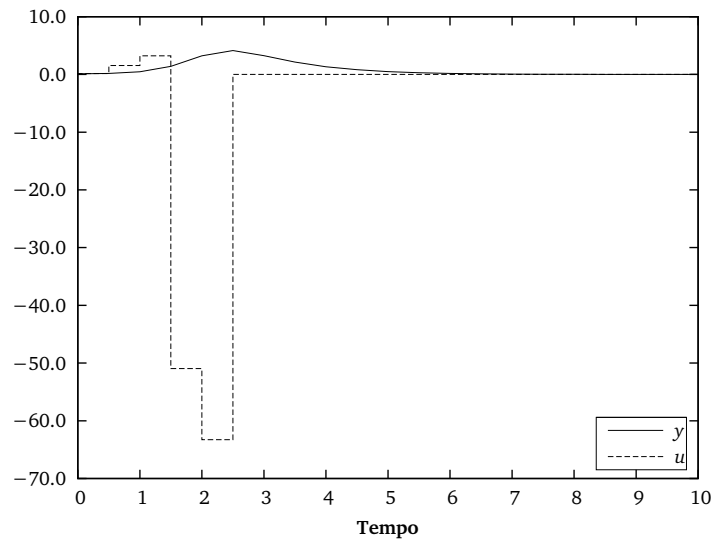
Ainda que sejam utilizadas outras factorizações, a utilização de permutações na matriz  $M_k$  resultará sempre em *fill-in*. No entanto, e uma vez que as características numéricas da factorização  $QR$  com pivotagem se devem à forma como são particionadas as variáveis do problema, esta pivotagem poderá ser reordenada, mantendo essa divisão, de forma a conseguir uma redução do *fill-in*.

### 5.5.3 Trajectórias de controlo

Na Figura 5.8 são apresentados os perfis das variáveis medida e manipulada obtidos por resolução do problema de controlo com pivotagem, para um estado inicial  $x_0 = 0.1$ . Os parâmetros de controlo utilizados foram  $Q_y = Q_u = I$ , com um horizonte de previsão  $p = 30$  e um horizonte e controlo  $m = 10$ . Estes resultados coincidem com a solução analítica do mesmo problema.

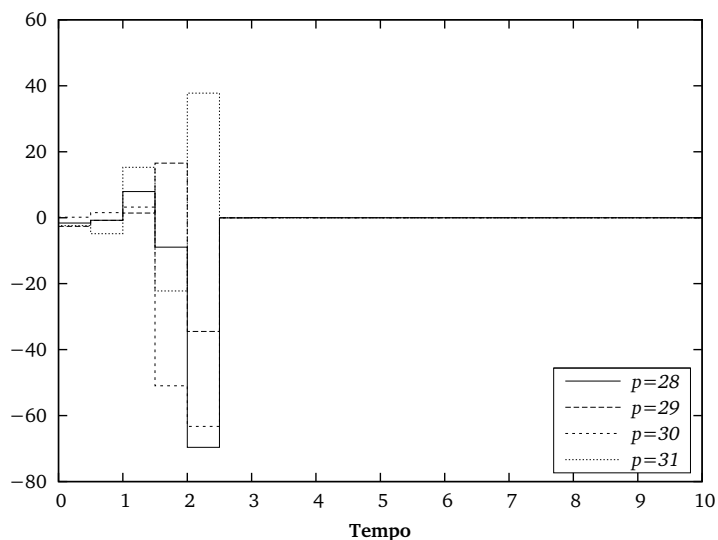
Na Figura 5.9 são apresentados os perfis obtidos, com os mesmos parâmetros de controlo, sem utilização de pivotagem.

Verifica-se que a solução obtida é totalmente distinta da solução analítica do problema, o que se deve ao mau condicionamento numérico do problema de optimização, como se pode verificar através do número de condição da matriz Hessiana reduzida apresentado na Tabela 5.3. A instabilidade numérica dos resultados obtidos sem pivotagem é confirmada pelos resultados apresentados na Figura 5.10, onde são apresentados os perfis de controlo obtidos para diferentes horizontes de previsão



**Figura 5.9** Perfil de controlo obtido utilizando a factorização  $QR$  sem pivotagem para o sistema LTI instável.

( $p = 28, 29, 30, 31$ ).



**Figura 5.10** Perfis de controlo obtidos, sem pivotagem, para o sistema LTI instável, com horizontes de controlo de tamanho  $p = 28, 29, 30$  e  $31$ .

Para qualquer um dos horizontes de previsão considerados, a solução numérica do problema é distinta da solução analítica, sendo também diferentes as soluções obtidas para diferentes horizontes de previsão. Estas diferenças devem-se aos diferentes erros numéricos na solução de cada um dos problemas, devidos ao elevado número de condição da Hessiana projectada utilizada na solução do problema de optimização. Quando é utilizada a factorização com pivotagem, os perfis obtidos para estes horizontes coincidem entre si e com a solução analítica.

Em alternativa à utilização da factorização QR foi considerada a resolução do problema quadrático utilizando a rotina SQOPT (Gill et al., 1997b). Com esta rotina não é possível obter uma solução para o problema quadrático correspondente ao problema de controlo do sistema LTI instável com um horizonte de previsão,  $p$ , superior a 23. A rotina não encontra uma solução, e dá a informação de que o problema é ilimitado ou *badly scaled*. Utilizando a rotina DGEEQU, da biblioteca LAPACK, foram calculados factores de escala para as linhas e colunas da matriz  $M_k$ . Os resultados desta rotina indicam que todos os elementos desta matriz têm magnitudes semelhantes, não sendo necessária a aplicação dos factores de escala calculados. Uma vez que é obtida uma solução para este problema utilizando a partição das variáveis resultante da factorização QR com pivotagem, a falha na SQOPT será devida à partição das variáveis escolhida para a decomposição da solução. Este exemplo demonstra como é crucial, para a robustez da solução do problema de controlo, uma partição adequada das variáveis do problema.

## 5.6 Aplicação ao processo TE

A reformulação proposta, utilizando a decomposição QR com pivotagem, demonstra que uma selecção criteriosa das variáveis independentes do problema permite resolver problemas de optimização dinâmica de sistemas instáveis utilizando métodos sequenciais. A utilização da factorização QR ainda que robusta numericamente, não é particularmente eficiente do ponto de vista computacional. Uma das desvantagens da sua utilização é o facto de não ser possível tirar partido da estrutura e esparsidade dos problemas, o que tornará ineficiente a sua aplicação a problemas de grande dimensão. Por esse motivo, e para aplicação ao problema de controlo do processo TE, foram procuradas alternativas que, resolvendo problemas quadráticos por decomposição do passo, permitissem tirar partido da estrutura do problema e, ao mesmo tempo, seleccionar as variáveis independentes do problema de forma numericamente robusta.

Considerou-se, por isso, a utilização da rotina SQOPT (Gill et al., 1997b), concebida para a solução de problemas quadráticos com matrizes Hessianas e de restrições esparsas. Esta rotina implementa um método de decomposição da solução em que a partição das variáveis em dependentes e independentes é feita de forma a que a matriz do espaço nulo,  $Z$ , seja bem condicionada. Para tal é feita uma factorização LU, com pivotagem, da matriz das restrições utilizando a rotina LUSOL (Gill et al., 1987), que implementa um conjunto de regras destinadas a manter a esparsidade e a estabilidade numérica. Ainda que esta partição não seja tão robusta como a obtida com a factorização QR com pivotagem, revelou-se suficiente para resolver o problema de controlo do processo TE.

**Tabela 5.4** Variação do número de condição da Hessiana projectada do problema de controlo do processo TE ( $Q_y = Q_u = I$ ,  $m = 5$ ).

	$p = 30$	$p = 50$	$p = 70$
c/pivotagem	274.2	400.0	881.7
s/pivotagem	$4.0 \times 10^5$	$7.9 \times 10^6$	$1.4 \times 10^8$

**Tabela 5.5** Variação do número de condição da Hessiana projectada do problema de controlo do processo TE ( $Q_y = I$ ,  $Q_u = 0$ ,  $m = 5$ ).

	$p = 30$	$p = 50$	$p = 70$
c/pivotagem	2128.8	3164.6	6993.2
s/pivotagem	$3.2 \times 10^6$	$6.3 \times 10^7$	$1.1 \times 10^9$

Os resultados apresentados na Secção seguinte, relativos ao condicionamento numérico do problema foram obtidos utilizando factorização  $QR$  com e sem pivotagem. Estes resultados apenas pretendem demonstrar, novamente, os benefícios possíveis, do ponto de vista da robustez numérica, de uma partição adequada das variáveis do problema. Os resultados de controlo apresentados nas restantes Secções foram obtidos utilizando a rotina SQOPT para resolver os problemas quadráticos.

### 5.6.1 Condicionamento numérico

Nas Tabelas 5.4 e 5.5 é apresentado o número de condição da matriz Hessiana projectada para o problema de controlo do processo TE, utilizando diferentes parâmetros de controlo. São apresentados resultados para a utilização da factorização  $QR$  com e sem pivotagem.

Novamente se verifica, para os dois conjuntos de parâmetros de controlo, que o número de condição da matriz Hessiana projectada cresce de forma exponencial com o tamanho do horizonte de previsão, quando não é efectuada pivotagem. Este crescimento é evitado pela utilização de pivotagem, situação em que esse crescimento é aproximadamente linear. Verifica-se também que, para qualquer combinação de parâmetros de controlo e de dimensões dos horizontes, o número de condição da Hessiana projectada é sempre significativamente inferior quando é utilizada a pivotagem no particionamento das variáveis.

### 5.6.2 Aplicação da estratégia de horizontes adaptáveis

Nesta secção são apresentados os resultados da aplicação da estratégia de horizontes adaptáveis, descrita na Secção 5.3, ao problema de controlo do processo TE, na

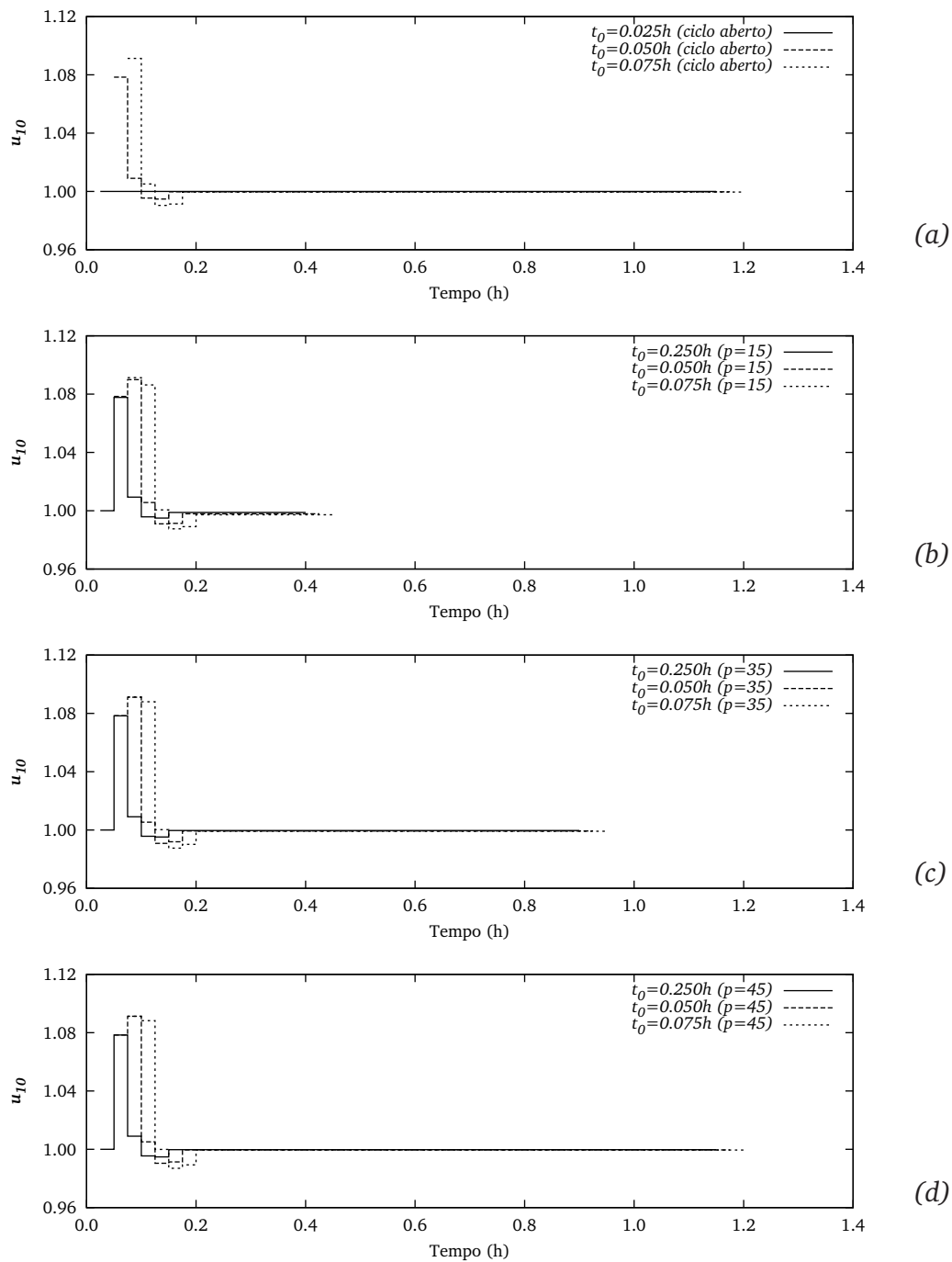
presença de uma perturbação em degrau de  $+5^{\circ}\text{C}$  na temperatura da água de arrefecimento do reactor, a partir do instante inicial da simulação. Os parâmetros utilizados no problema de controlo são  $p = 45$ ,  $m = 5$ ,  $T = 0.025\text{h}$ ,  $Q_u = 0$ ,  $Q_y = 0.1$ , para a generalidade das medidas,  $Q_y = 10$  para as composições de  $G$  e  $H$  no produto e para o caudal de produto.

No instante inicial de simulação,  $t_0 = 0\text{h}$ , o problema é facilmente resolvido, uma vez que a perturbação ainda não se fez sentir ( $x_0 = 1$ ), sendo as equações do modelo facilmente integradas em todo o horizonte de previsão. O perfil de controlo resultante da solução do problema é  $U = 1$ , utilizando variáveis normalizadas, para todo o horizonte. No instante de amostragem seguinte,  $t_0 = 0.025\text{h}$  a existência da perturbação leva a que já exista algum desvio relativamente ao estado estacionário inicial. Neste instante de amostragem, na primeira iteração, as equações do modelo são integradas utilizando uma trajectória nominal de controlo resultante da solução do problema no instante de amostragem anterior ( $\bar{U} = 1$ ). No caso da posição da válvula da água de arrefecimento do reactor, variável  $u_{10}$ , esta trajectória é representada na Figura 5.11a. Por integração das equações do modelo, utilizando esta trajectória de controlo, obtém-se o perfil apresentado na Figura 5.12a, para a pressão do reactor. Verifica-se que, no instante  $t = 0.8$ , a pressão do reactor ultrapassa o limite superior de segurança, resultando numa falha na integração das equações do modelo e das sensibilidades. Não é, por isso, possível integrar as equações do modelo ao longo de todo o horizonte de previsão ( $p = 45$ , correspondendo a um instante final igual a 1.15h).

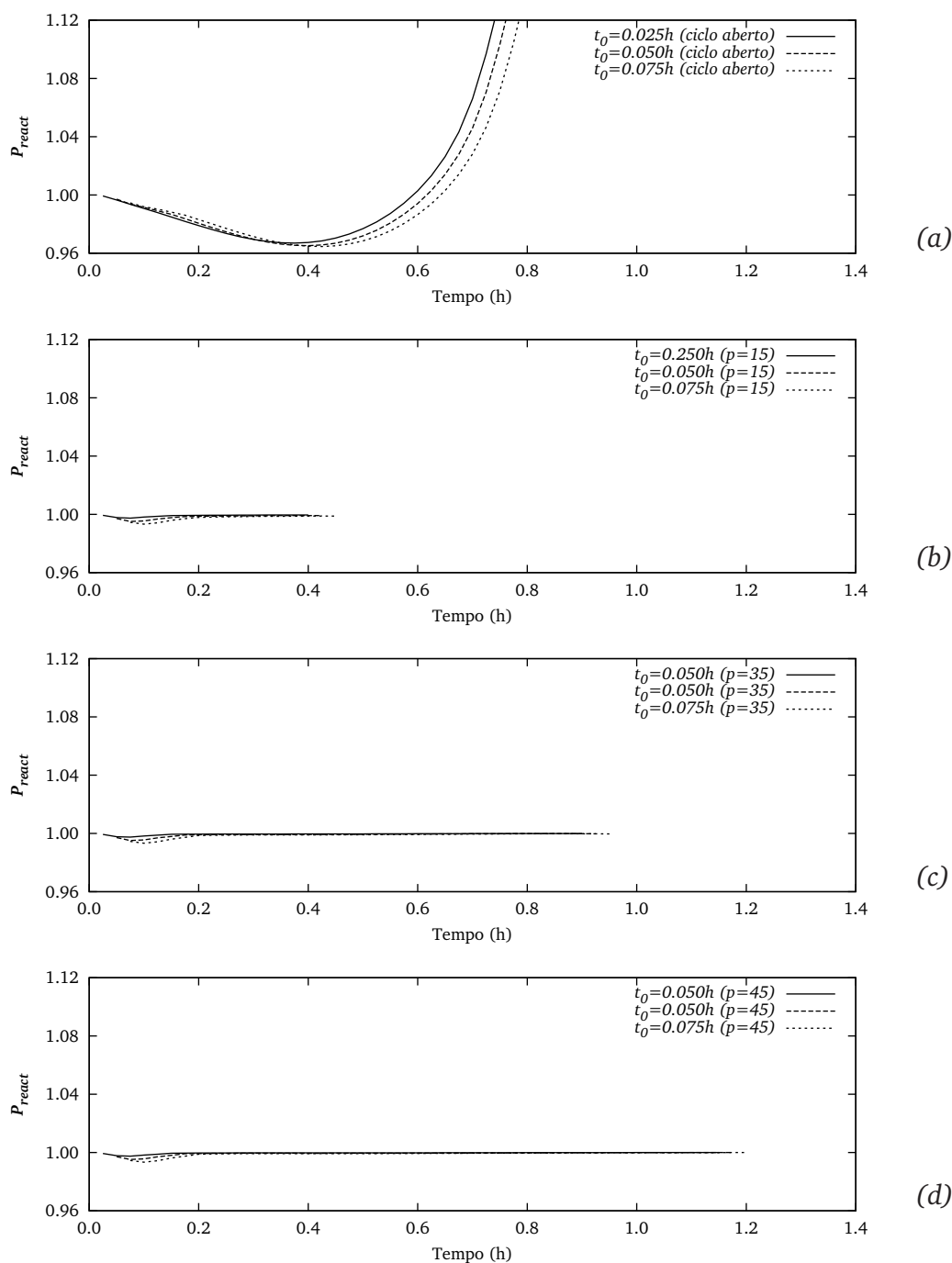
Para ultrapassar este problema, o tamanho do horizonte de previsão foi reduzido para  $p = 15$ , tornando possível a integração das equações do modelo em todo o horizonte de previsão. O problema de controlo é resolvido, sendo obtida a trajectória de  $u_{10}$  apresentada na Figura 5.11b e a trajectória da pressão apresentada na Figura 5.12b. A solução de controlo obtida com este horizonte é depois utilizada como solução inicial do problema de controlo com um horizonte de previsão  $p = 35$ . Na Figura 5.12c pode verificar-se que a trajectória de controlo obtida já é suficiente para estabilizar o processo além do instante em que inicialmente ocorreu a falha. A solução deste problema é utilizada como aproximação inicial ao problema de controlo original, com  $p = 45$ , sendo já possível integrar as equações ao longo de todo o horizonte de previsão. A solução deste problema é apresentada nas Figuras 5.11d e 5.12d.

O mesmo problema se verifica para os problemas de controlo com instantes iniciais  $t_0 = 0.05\text{h}$  e  $t_0 = 0.075\text{h}$ , o que se deve ao facto de as soluções de controlo serem obtidas sem que o controlador tenha conhecimento da existência da perturbação. Em cada um destes instantes é possível resolver o problema de controlo com  $p = 45$ , ainda que em todos se verifique uma falha na integração das equações do modelo com a trajectória nominal inicial.

É de notar que o tamanho dos horizontes escolhidos privilegia sobretudo a robustez, e



**Figura 5.11** Perfis da posição da válvula da água de arrefecimento do reactor, variável manipulada  $u_{10}$ , nos instantes de amostragem 0.025h, 0.05h e 0.075h; a) perfis nominais iniciais; b) perfis óptimos para  $p = 15$ ; c) perfis óptimos para  $p = 35$ ; d) perfis óptimos para  $p = 45$ .



**Figura 5.12** Previsões da pressão do reactor nos instantes de amostragem 0.025h, 0.05h e 0.075h, obtidos por integração do modelo, utilizando diferentes perfis de variáveis de controlo; a) perfis nominais iniciais; b) perfis óptimos para  $p = 15$ ; c) perfis óptimos para  $p = 35$ ; d) perfis óptimos para  $p = 45$ .

não tanto a eficiência computacional. Ainda assim, o número de iterações necessárias é elevado sobretudo quando as trajectórias nominais ainda se encontram afastadas da solução óptima, o que ocorre nos problemas com horizontes de previsão mais curtos. Nos problemas com horizontes maiores, a solução inicial já é, geralmente, uma boa aproximação da solução óptima, pelo que o problema converge rapidamente. Ou seja, ainda que o número de iterações total necessário com este tipo de estratégia seja grande, muitas destas iterações correspondem a problemas de pequena dimensão. No caso dos resultados apresentados são necessárias duas a três iterações para a solução do problema com  $p = 15$ . As soluções dos problemas com  $p = 35$  e  $p = 45$  são geralmente obtidas em uma ou duas iterações.

Deve ainda ser referido que os horizontes de controlo utilizados foram escolhidos de forma totalmente heurística, sendo possível equacionar a utilização de outros conjuntos de horizontes permitindo igualmente a solução do problema.

### 5.6.3 Resultados de controlo

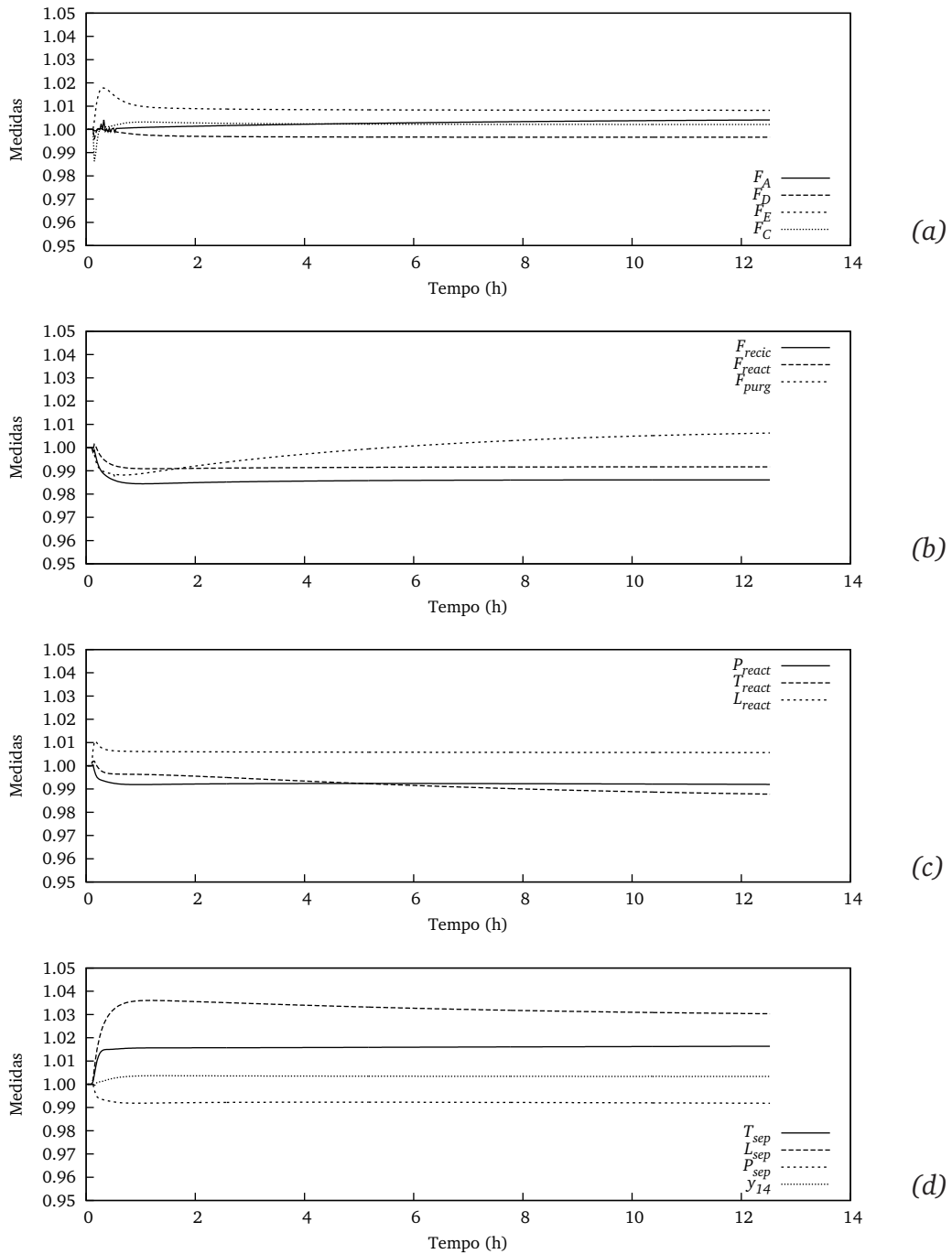
Para testar o desempenho do algoritmo de controlo reformulado, este foi utilizado para resolver o problema de controlo do processo TE. Como referido na Secção 4.3, devido à instabilidade em ciclo aberto e ao mau condicionamento numérico, o algoritmo apresentado na Secção 4.1 não é capaz de, por si só, resolver este problema. Para que fosse possível a resolução do problema utilizando este algoritmo foram necessárias alterações na estrutura do processo, nomeadamente a utilização de um compensador interno.

Usando a formulação de controlo com as modificações descritas, na Figura 5.13 e 5.13 são apresentados, respectivamente, os perfis das variáveis medidas e manipuladas obtidos com a ocorrência de um degrau de  $+5^{\circ}\text{C}$  na temperatura da água de arrefecimento do reactor a partir do instante  $t = 0.1\text{h}$ . Os parâmetros de controlo utilizados para obter estes perfis foram  $p = 45$ ,  $m = 5$ ,  $T = 0.025\text{h}$ ,  $Q_y = 0.1$  para a generalidade das variáveis medidas,  $Q_y = 10$  para as composições de  $G$  e  $H$  no produto ( $y_{40}$  e  $y_{41}$ ) e para o caudal de produto ( $y_{17}$ ), e  $Q_u = 0$  para todas as variáveis manipuladas.

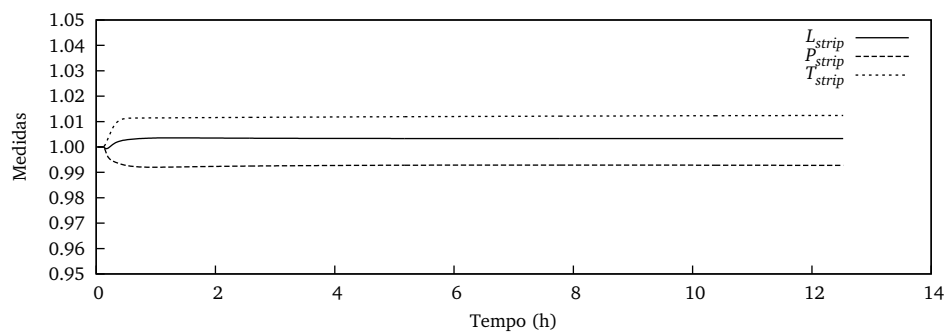
Verifica-se que o controlador é capaz de rejeitar rapidamente a perturbação. As variações observadas em todas as variáveis medidas são pequenas, mantendo-se a pressão do reactor e os níveis do reactor, separador e *stripper* afastados dos seus limites de segurança. As variações nas composições de  $G$  e  $H$  no produto e o seu caudal apresentam variações inferiores a 1%. Verifica-se a existência de algum *off set* que poderia ser eliminado pela utilização de acção integral.

Este resultados demonstram que as alterações propostas neste Capítulo ao algoritmo de controlo predictivo não-linear apresentado na Secção 4.1 permitem agora a sua

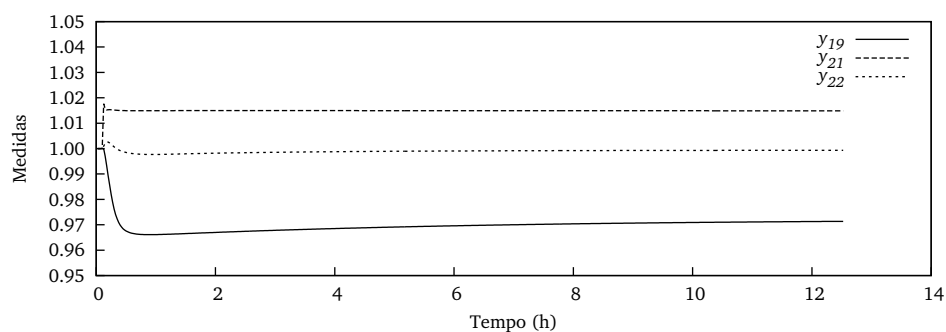




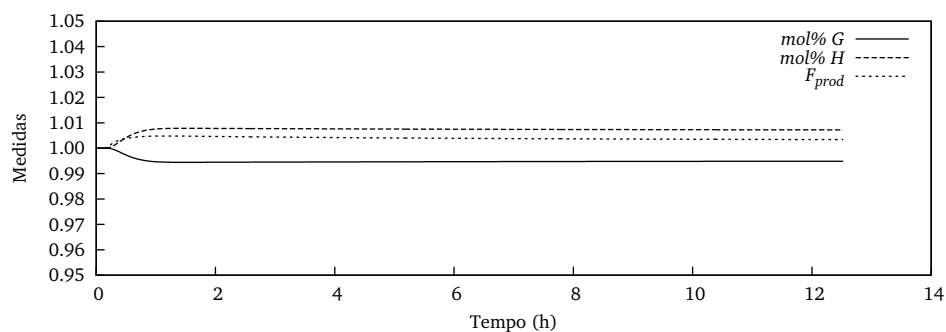
**Figura 5.13** Perfis das variáveis medidas para uma perturbação em degrau de +5°C na temperatura de entrada da água de arrefecimento do reactor.



(e)

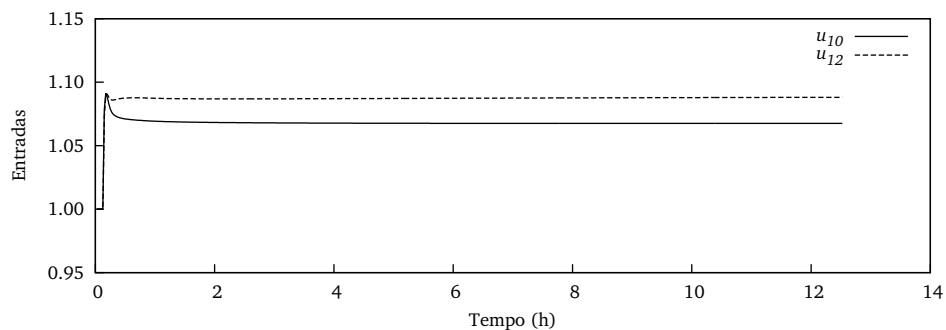


(f)



(g)

**Figura 5.13 (cont.)** Perfis das variáveis medidas para uma perturbação em degrau de  $+5^{\circ}\text{C}$  na temperatura de entrada da água de arrefecimento do reactor.



**Figura 5.14** Perfis das variáveis manipuláveis para uma perturbação em degrau de  $+5^{\circ}\text{C}$  na temperatura de entrada da água de arrefecimento do reactor.

aplicação directa ao processo TE, instável em ciclo aberto e numericamente mal condicionado.

## 5.7 Conclusões

Neste Capítulo foi apresentada uma estratégia para lidar com o problema de mau condicionamento numérico surgido na aplicação do algoritmo de controlo predictivo proposto por Oliveira (1994), descrito na Secção 4.1. Esta estratégia consiste na utilização de uma representação diferente das equações do modelo linearizado no horizonte de previsão, que permite a utilização de decomposições robustas, como a factorização *QR* com pivotagem. A utilização desta factorização, juntamente com a solução do problema quadrático utilizando um método de decomposição do passo, permite a obtenção de problemas de optimização numericamente bem condicionados. No âmbito da solução do problema de controlo predictivo não-linear (4.2) a abordagem proposta neste Capítulo é, do ponto de vista global, equivalente à da formulação de controlo de Oliveira (1994), correspondendo a uma aproximação de Gauss-Newton. A diferença entre as duas formulações reside na forma como os problemas quadráticos são resolvidos.

Os resultados obtidos mostram que a melhoria no condicionamento numérico do problema é conseguido pela utilização da pivotagem que, ao nível do problema de optimização, pode ser interpretada como a escolha dos conjuntos de variáveis dependentes e independentes. Na estratégia proposta, ao contrário do algoritmo anterior, as variáveis independentes do problema podem ser qualquer combinação de variáveis manipuláveis, de estado e de saída, sendo a escolha feita de forma a conseguir um problema bem condicionado. Os resultados apresentados mostram que esta escolha está também relacionada com a estabilidade do sistema. Por exemplo, no caso do sistema LTI instável utilizado, o algoritmo identifica o modo instável, cujo valor no último instante do horizonte de previsão passa a ser estabelecido pelo algoritmo de optimização. Podem ser estabelecidas analogias entre esta escolha e a utilização de restrições terminais. Também pode ser estabelecida uma analogia entre a utilização de qualquer variável de saída, ou de estado, como variável independente e a introdução de compensadores internos entre o algoritmo de controlo predictivo e o processo. A partição das variáveis do problema é feita de forma sistemática, utilizando algoritmos bem estabelecidos, ao contrário da escolha dos compensadores que, geralmente, é feita de forma não sistemática, tal como é possível verificar na análise feita na Secção 3.2 às estruturas de controlo propostas por outros autores para o processo TE.

A importância da forma como são particionadas as variáveis do problema é revelada pelo exemplo do sistema LTI instável. Para este sistema a rotina SQOPT não é capaz de resolver o problema de controlo para horizontes de previsão com tamanho superior

a  $p = 30$ . Já a utilização da factorização QR com pivotagem permite resolver este problema.

A abordagem proposta depende da integração, em ciclo aberto, das equações do modelo no horizonte de previsão, sendo possível a ocorrência de falhas, sobretudo no caso de sistemas instáveis. Para evitar estas falhas é utilizada uma estratégia de horizontes de previsão ajustáveis, adaptada da estratégia *incremental single shooting* apresentada por [Michalik et al. \(2009\)](#).

A formulação de controlo predictivo resultante das estratégias propostas para lidar com o mau condicionamento e com a instabilidade em ciclo aberto é aplicada, com sucesso, ao problema de controlo predictivo do processo TE, sem necessidade de utilizar camadas adicionais de controlo.

Ainda que a estratégia proposta permita melhorar a robustez dos métodos sequenciais, não deve ser considerada suficiente para garantir sempre a obtenção de soluções de controlo óptimo com sistemas instáveis em ciclo aberto, uma vez que continua a depender da integração das equações do modelo em todo o horizonte. Como já referido, os métodos actuais que oferecem maiores garantias relativamente à eliminação desta dificuldades são os métodos simultâneos, ainda que também possa ser necessário o recurso a decomposições robustas como a apresentada na Secção 5.2.

A utilização da factorização QR na solução do problema em todos os instantes de amostragem pode tornar esta abordagem demasiado dispendiosa, do ponto de vista do esforço computacional. No entanto, esta factorização não é estritamente necessária, uma vez que a sua característica que permite estabilizar o desempenho numérico do algoritmo é a matriz de permutações utilizada. Assim, a mesma divisão de variáveis poderá ser utilizada na solução do problema em múltiplos instantes de amostragem consecutivos, utilizando outras factorizações para resolver os sistemas lineares. Estas factorizações poderão ser feitas tirando partido da estrutura do problema. Será também conveniente que a factorização utilizada permita obter uma estimativa do número de condição da matriz, que poderá ser utilizado para decidir se é, ou não, necessário obter uma nova partição das variáveis.

# Capítulo 6

## Conclusões globais

O trabalho apresentado nesta tese centrou-se na aplicação de métodos de optimização dinâmica, mais especificamente de controlo predictivo não-linear (NMPC) a processos químicos. Partindo do algoritmo proposto por [Oliveira \(1994\)](#), que implementa um método sequencial, foi proposta uma nova formulação de controlo. Com esta formulação, também de natureza sequencial, pretenderam-se ultrapassar as dificuldades sentidas na aplicação directa de métodos desta classe a sistemas que sejam instáveis em ciclo aberto, ou que resultem em problemas de optimização mal condicionados numericamente.

Neste Capítulo são apresentadas as principais conclusões obtidas a partir do trabalho efectuado, bem como as suas principais contribuições. São também apresentadas algumas sugestões para trabalho futuro.

As principais contribuições do presente trabalho podem dividir-se em três categorias, a caracterização do processo Tennessee Eastman (Capítulo 3), a aplicação do algoritmo de controlo predictivo proposto por [Oliveira \(1994\)](#) a este processo (Capítulo 4), e a reformulação deste mesmo algoritmo para aplicação a sistemas instáveis em ciclo aberto e a problemas numericamente mal condicionados (Capítulo 5).

### **Caracterização do processo Tennessee Eastman**

Utilizando o modelo do processo foi feita a sua caracterização com o objectivo de avaliar a pertinência da aplicação de controlo predictivo não-linear a este processo. Os resultados obtidos demonstram que o processo é instável em ciclo aberto e que possui dois modos integrantes. É também possível verificar a forte interacção existente entre as diferentes unidades. A determinação das condições de operação óptimas, do ponto de vista do custo e da capacidade de produção, permitem determinar os conjuntos de restrições activas em cada uma destas condições, que se verificam ser diferentes e dependentes da composição pretendida para o produto. Os resultados obtidos indicam

que é expectável um melhor desempenho de uma estratégia de controlo baseada em modelos, relativamente à utilização de estruturas de controlo descentralizadas, como as propostas por diversos autores.

A análise feita permite ainda aferir a importância da existência de um modelo dinâmico detalhado do processo. Este modelo permite adquirir uma compreensão sobre o comportamento e as características do processo que dificilmente seriam obtidas a partir da sua operação.

### **Controlo predictivo do processo TE**

Em face dos resultados obtidos pela caracterização do processo, foi decidido aplicar a formulação de NMPC proposta por Oliveira (1994) inicialmente ao reactor e depois a todo o processo. A aplicação directa desta formulação revelou-se inviável, em ambos os casos, devido ao mau condicionamento numérico do problema de optimização, resultante da instabilidade. A forma encontrada de lidar com este problema foi a utilização de um compensador. No caso do problema de controlo do reactor, o compensador foi escolhido e sintonizado de forma a conseguir a estabilização antes da aplicação do algoritmo de NMPC. No caso do processo, verificou-se que o compensador utilizado no problema de controlo do reactor, ainda que não suficiente para estabilizar o processo, permite melhorar significativamente o condicionamento numérico do problema de controlo, sendo esta melhoria suficiente para permitir a sua solução. Os resultados de controlo obtido mostram que, de facto, a utilização de NMPC é a solução de controlo mais adequada para problemas com características semelhantes às do processo TE.

A utilização de uma estrutura de controlo em cascata, quando a escolha do(s) compensador(es) seja simples, será uma solução de interesse, pela sua simplicidade, exequibilidade, e por permitir resolver o problema. No entanto, deve ser considerado o facto de, na generalidade dos processos não-lineares de média ou grande dimensão, a sua escolha não ser possível de forma sistemática, sendo necessário o recurso a regras heurísticas, que poderão resultar em soluções sub-óptimas.

### **Reformulação da estratégia de controlo**

Os resultados obtidos revelam as principais dificuldades na aplicação de métodos sequenciais a sistemas instáveis em ciclo aberto. Para que este tipo de métodos possa ser aplicado directamente a estes sistemas, foram introduzidas algumas modificações no algoritmo de controlo, motivadas pelos resultados obtidos com a utilização de um compensador. Estas modificações ocorreram ao nível da solução do problema quadrático, nomeadamente na forma como são utilizadas as equações do modelo e as suas variáveis. Ao contrário da formulação anterior, em que apenas as variáveis manipuladas são variáveis de decisão, com a nova formulação passa ser possível a

utilização de diferentes combinações de qualquer tipo de variáveis do modelo (de estado, manipuladas e medidas). As variáveis do modelo são particionadas em variáveis independentes, utilizadas como variáveis de decisão do problema de optimização, e variáveis independentes, determinadas pelas equações do modelo. Esta partição é depois utilizada no problema quadrático, resolvido por decomposição da solução. Desta forma, é possível manter a pequena dimensão dos problemas de optimização, uma das características desejáveis dos métodos sequenciais

A forma como é feita a partição das variáveis é crucial para o sucesso desta formulação, como demonstram os resultados apresentados. Para garantir que as variáveis são particionadas de forma a obter um problema de optimização bem condicionado numericamente, é utilizada a factorização  $QR$ , com pivotagem, da matriz que representa as equações linearizadas do modelo em todo o horizonte de previsão. A pivotagem das colunas determina a partição das variáveis do problema, sendo a robustez numérica da formulação proposta devida a esta pivotagem.

Esta abordagem ao problema de controlo depende da integração em ciclo aberto das equações do modelo e das sensibilidades, nomeadamente para obtenção das derivadas de primeira ordem necessárias para a linearização do modelo. Para evitar as consequências de falhas nesta integração, é utilizada uma estratégia de horizontes de previsão ajustáveis, adaptada da estratégia apresentada por [Michalik et al. \(2009\)](#), que considera a utilização de horizontes de previsão sucessivamente crescentes. A utilização sucessiva da solução de controlo obtida em horizontes mais pequenos permite uma estabilização que poderá ser suficiente para a integração das equações do modelo num horizontes progressivamente mais longos, até que seja possível a obtenção da solução de controlo no horizonte de previsão com a dimensão pretendida.

A capacidade da formulação proposta em resolver directamente problemas de NMPC de sistemas instáveis em ciclo aberto é demonstrada pela sua aplicação, com sucesso, ao processo TE.

## **6.1 Perspectivas de trabalho futuro**

Para que a formulação proposta neste trabalho possa ser mais ampla, devem ser consideradas ainda vários desenvolvimentos. Alguns destes são apresentados de seguida.

### **Implementação da formulação de controlo proposta**

Ainda que a formulação de controlo resultante das utilizações das estratégias descritas no Capítulo 5 já se encontre implementada computacionalmente (em FORTRAN), esta implementação deve ainda ser objecto de melhorias, quer ao nível de robustez, quer ao nível da eficiência. Ao nível da robustez os aspectos merecedores de maior atenção

serão uma melhor implementação dos métodos de pesquisa linear e a resolução do problema na presença de restrições activas. Ao nível da eficiência computacional, os aspectos mais relevantes a considerar serão a capacidade de realizar a factorização QR sem formação explícita da matriz e dos seus factores, tirando partido da sua estrutura e esparsidade, a possibilidade de, em cada iteração, actualizar os factores já calculados previamente e a utilização de factorizações mais eficientes (Li e Saad, 2006). Outro aspecto a ter em consideração será, como já referido, a reutilização da mesma pivotagem em diferentes instantes de amostragem.

### **Estimativa de estados**

Neste trabalho não foi considerado o problema da estimativa de estados, isto é, da obtenção das variáveis de estado do processo a partir dos valores das variáveis medidas e de entrada no instante actual e, eventualmente, num horizonte temporal passado. A resolução deste problema é estritamente necessária para a implementação de qualquer algoritmo de MPC, uma vez que é a utilização do estado do processo no instante de amostragem em que o problema é resolvido que confere a estes controladores características de controlo por realimentação. Neste âmbito podem ser utilizadas várias ferramentas, como filtros de Kalman (Grewal e Andrews, 2001) ou estimação em horizontes móveis (Rao, 2000; Rao e Rawlings, 2002). Estes últimos são particularmente úteis no caso de processos não-lineares, sendo o problema resolvido semelhante ao problema de NMPC. Pode ainda ser considerada a utilização de *particle filtering* (Rawlings e Mayne, 2009), aplicável a processos não-lineares. Este método é particularmente útil no caso em que o ruído aleatório que afecta as variáveis do processo não pode ser considerado normal. Pode ser aplicado isoladamente, ou em combinação com a estimação em horizontes móveis.

### **Construção de modelos**

Os resultados apresentados neste trabalho, relativos à caracterização do processo TE e ao seu controlo, demonstram a importância da existência de um modelo do processo capaz de descrever o seu comportamento numa gama alargada de condições de operação. Um modelo com estas características deve ser baseado em princípios fundamentais e será, geralmente, não-linear. A complexidade destes modelos poderá ser diminuída se for possível a identificação das relações existentes entre as diferentes variáveis do processo. Neste âmbito, deve ainda ser considerado o problema da estimativa de parâmetros, que poderá ser feita fora de linha, ou então em simultâneo com a estimativa de estados, utilizando os métodos de estimação com horizontes móveis atrás descritos.

### **Linguagens para implementação de algoritmos de optimização dinâmica**

Será de todo o interesse a implementação da estratégia de NMPC em linguagens mais



flexíveis, que tornem mais simples a utilização de diferentes abordagens e diferentes *solvers* para um mesmo problema. Neste âmbito, será particularmente relevante a ferramenta TACO (Kirches e Leyffer, 2011), uma extensão da linguagem AMPL (Fourer et al., 1990) para resolução de problemas de controlo óptimo. Ainda que menos flexível, mas com a vantagem de ser *open source*, poderá também ser considerada a utilização da ferramenta OPTIMICA (Åkesson et al., 2010), uma extensão da linguagem MODELICA para resolução de problemas de optimização dinâmica.

### **Incorporação de incertezas**

Neste trabalho não foi considerada a existência de incertezas, sendo todas as variáveis tratadas como quantidades determinísticas. A existência de incertezas, a sua quantificação e a sua propagação no futuro, no âmbito da aplicação de MPC, foi já considerada noutros trabalhos (Gândara et al., 2006). Estas contribuições poderão ser integradas na nova formulação de NMPC para, por exemplo, determinar políticas de amostragem com custo mínimo.



# Bibliografia

- Åkesson, J., Årzén, K.-E., Gäfvert, M., Bergdahl, T., e Tummescheit, H. (2010). Modeling and optimization with Optimica and JModelica.org—Languages and tools for solving large-scale dynamic optimization problems. *Computers & Chemical Engineering*, 34(11):1737–1749.
- Albuquerque, J., Gopal, V., Status, G., Biegler, L. T., e Ydstie, B. E. (1999). Interior point SQP strategies for large-scale, structured process optimization problems. *Computers & Chemical Engineering*, 23(4):543–554.
- Anderson, E., Bai, Z., Bischof, C. H., Blackford, S., Demmel, J., Dongarra, J. J., Croz, J. D., Greenbaum, A., Hammarling, S., McKenney, A., e Sorensen, D. (1999). *LAPACK Users' Guide*. SIAM, Philadelphia, 3rd edition.
- Antelo, L. T., Banga, J. R., e Alonso, A. A. (2008). Hierarchical design of decentralized control structures for the Tennessee Eastman Process. *Computers & Chemical Engineering*, 32(9):1995–2015.
- Ascher, U. M. e Petzold, L. R. (1998). *Computer Methods for Ordinary Differential Equations and Differential–Algebraic Equations*. SIAM, Philadelphia.
- Balakrishna, S. e Biegler, L. T. (1992). Targeting strategies for the synthesis and energy integration of nonisothermal reactor networks. *Industrial & Engineering Chemistry Research*, 31(9):2152–2164.
- Banerjee, A. e Arkun, Y. (1995). Control configuration design applied to the Tennessee Eastman plant-wide control problem. *Computers & Chemical Engineering*, 19(4):453–480.
- Bartl, M., Li, P., e Biegler, L. T. (2011). Improvement of state profile accuracy in nonlinear dynamic optimization with the quasi-sequential approach. *AIChE Journal*, 57(8):2185–2197.
- Becerra, V., Abu-el zeet, Z., e Roberts, P. (1999). Integrating predictive control and economic optimisation. *Computing & Control Engineering Journal*, 10(5):198–208.
- Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.
- Benson, H. e Shanno, D. (2008). Interior-point methods for nonconvex nonlinear programming: regularization and warmstarts. *Computational Optimization and Ap-*

- plications*, 40(2):143–189.
- Biegler, L. T. (1984). Solution of Dynamic Optimization Problems by Successive Quadratic Programming and Orthogonal Collocation. *Computers & Chemical Engineering*, 8(3):243–248.
- Biegler, L. T. (1995). Convergence analysis for a multiplier-free reduced Hessian method. Technical Report EDRC 06-203-95, Engineering Design Research Center, Carnegie Mellon University, Pittsburgh, PA.
- Biegler, L. T. (2007). An overview of simultaneous strategies for dynamic optimization. *Chemical Engineering and Processing: Process Intensification*, 46(11):1043–1053.
- Biegler, L. T. (2010). *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*. SIAM, Philadelphia.
- Biegler, L. T., Cervantes, A. M., e Wächter, A. (2002). Advances in simultaneous strategies for dynamic process optimization. *Chemical Engineering Science*, 57(4):575–593.
- Biegler, L. T. e Grossmann, I. E. (2004). Retrospective on Optimization. *Computers & Chemical Engineering*, 28(8):1169–1192.
- Biegler, L. T., Nocedal, J., e Schmid, C. (1995). A Reduced Hessian Method for Large-Scale Constrained Optimization. *SIAM J. Optimization*, 5(2):314–347.
- Binder, T., Blank, L., Bock, H. G., Burlisch, R., Dahmen, W., Diehl, M., Kronseder, T., Marquardt, W., Schlöder, J. P., e von Stryk, O. (2001). Introduction to model based optimization of chemical processes on moving horizons. In Grötschel, M., Krumke, S. O., e Rambau, J., editors, *Online Optimization of largescale Systems: State of the Art*, pages 295–339. Springer.
- Bischof, C. H. e Quintana-Ortí, G. (1998). Computing Rank-Revealing QR Factorizations of Dense Matrices. *ACM Transactions on Mathematical Software*, 24(2):226–253.
- Blanco, A. M., Bandoni, J. A., e Biegler, L. T. (2004). Re-design of the Tennessee Eastman Challenge Process: An Eigenvalue Optimization Approach. In Floudas, A. e Agrawal, R., editors, *Foundations of Computer Aided Process Design*, pages 517–520, Austin. CACHE Corporation.
- Bock, H. G., Diehl, M., Kostina, E. A., e Schlöder, J. P. (2006). Constrained Optimal Feedback Control of Systems Governed by Large Differential Algebraic Equations. In Biegler, L., Ghattas, O., Heinkenschloss, M., Keyes, D., e van Bloemen Waanders, B., editors, *Real-Time and Online PDE-Constrained Optimization*. SIAM.
- Bock, H. G. e Plitt, K. J. (1984). A multiple shooting algorithm for direct solution of optimal control problems. In *Proceedings 9th IFAC World Congress Budapest*, pages 243–247. Pergamon Press.
- Brenan, K. E., Campbell, S. L., e Petzold, L. R. (1989). *Numerical Solution of Initial-*

- Value Problems in Differential Algebraic Equations*. North–Holland, New York.
- Brown, P. N., Hindmarsh, A. C., e Petzold, L. R. (1998). Consistent Initial Condition Calculation for Differential-Algebraic Systems. *SIAM Journal on Scientific Computing*, 19(5):1495–1512.
- Bryson, A. E. e Ho, Y. (1969). *Applied Optimal Control*. Ginn and Company, Massachusetts.
- Byrd, R. H., Hribar, M. E., e Nocedal, J. (1999). An Interior Point Algorithm for Large-Scale Nonlinear Programming. *SIAM Journal on Optimization*, 9(4):877–900.
- Byrd, R. H., Nocedal, J., e Waltz, R. A. (2006). KNITRO: An Integrated Package for Nonlinear Optimization. In di Pillo, G. e Roma, M., editors, *Large-Scale Nonlinear Optimization*, pages 35–59. Springer-Verlag.
- Camacho, E. F. e Bordons, C. (1998). *Model Predictive Control*. Springer-Verlag, London.
- Camacho, E. F. e Bordons, C. (2007). Nonlinear Model Predictive Control: An Introductory Review. In Findeisen, R., Allgöwer, F., e Biegler, L. T., editors, *Assessment and Future Directions of Nonlinear Model Predictive Control*, volume 358 of *Lecture Notes in Control and Information Sciences*, pages 1–16. Springer.
- Cao, Y., Li, S., Petzold, L., e Serban, R. (2003). Adjoint Sensitivity Analysis for Differential-Algebraic Equations: The Adjoint DAE System and Its Numerical Solution. *SIAM Journal on Scientific Computing*, 24(3):1076–1089.
- Caracotsios, M. e Stewart, W. E. (1985). Sensitivity Analysis of Initial Value Problems with Mixed ODEs and Algebraic Equations. *Computers & Chemical Engineering*, 9(4):359–365.
- Cervantes, A. M. e Biegler, L. T. (2000). A stable elemental decomposition for dynamic process optimization. *Journal of Computational and Applied Mathematics*, 120(1–2):41–57.
- Cervantes, A. M. e Biegler, L. T. (2009). Optimization strategies for dynamic systems. In Floudas, C. A. e Pardalos, P. M., editors, *Encyclopedia of Optimization*, pages 2847–2858. Springer, 2nd edition.
- Cervantes, A. M., Wächter, A., Tütüncü, R. H., e Biegler, L. T. (2000). A reduced space interior point strategy for optimization of differential algebraic systems. *Computers & Chemical Engineering*, 24(1):39–51.
- Cuthrell, J. E. e Biegler, L. T. (1987). On the Optimization of Differential–Algebraic Process System. *AIChE Journal*, 33(8):1257–1270.
- Dennis, J. E., Heikenschloss, M., e Vicente, L. N. (1994). Trust–Region Interior–Point SQP Algorithms for a Class of Nonlinear Programming Problems. Technical report TR94–45, Department of Computational and Applied Mathematics, Rice University.
- Dennis, J. E. e Schnabel, R. B. (1983). *Numerical Methods for Unconstrained Optimi-*

- zation and Nonlinear Equations. Prentice–Hall, New Jersey.
- Deuffhard, P., Hairer, E., e Zugck, J. (1987). One-step and extrapolation methods for differential-algebraic systems. *Numerische Mathematik*, 51(5):501–516.
- Diehl, M., Bock, H. G., e Schlöder, J. P. (2005). A real-time iteration scheme for nonlinear optimization in optimal feedback control. *SIAM J. Control Optim.*, 43(5):1714–1736.
- Downs, J. J. e Vogel, E. F. (1993). A Plant–Wide Industrial Process Control Problem. *Computers & Chemical Engineering*, 17(3):245–255.
- Drud, A. S. (1994). CONOPT–A Large-Scale GRG Code. *INFORMS Journal on Computing*, 6(2):207–216.
- Duff, I. S. e Reid, J. K. (1993). MA48, a Fortran code for direct solution of sparse unsymmetric linear systems of equations. Technical Report RAL–93–072, Rutherford Appleton Laboratory.
- Duvall, P. M. e Riggs, J. B. (2000). On–line Optimization of the Tennessee Eastman Challenge Problem. *Journal Of Process Control*, 10(1):19–33.
- Esposito, W. R. (2009). Hamilton-Jacobi-Bellman Equation. In Floudas, C. A. e Pardalos, P. M., editors, *Encyclopedia of Optimization*, pages 1473–1476. Springer, 2nd edition.
- Feehery, W. F. e Barton, P. I. (1998). Dynamic Optimization With State Variable Path Constraints. *Computers & Chemical Engineering*, 22(9):1241–1256.
- Feehery, W. F., Tolsma, J. E., e Barton, P. I. (1997). Efficient sensitivity analysis of large-scale differential-algebraic systems. *Applied Numerical Mathematics*, 25(1):41–54.
- Fletcher, R. (1987). *Practical Methods of Optimization*. John Wiley & Sons, 2nd edition.
- Fletcher, R., Gould, N. I. M., Leyffer, S., Toint, P. L., e Wächter, A. (2002). Global Convergence of a Trust-Region SQP-Filter Algorithm for General Nonlinear Programming. *SIAM Journal on Optimization*, 13(3):635–659.
- Fletcher, R. e Leyffer, S. (2002). Nonlinear programming without a penalty function. *Mathematical Programming*, 91(2):239–269.
- Fletcher, R., Leyffer, S., e Toint, P. (2007). A brief history of filter methods. *SIAG/OPT Views-and-News, A Forum for the SIAM Activity Group on Optimization*, 1(1):2–12.
- Fourer, R., Gay, D. M., e Kernighan, B. W. (1990). A Modeling Language for Mathematical Programming. *Management Science*, 36:519–554.
- Gândara, J. F. M., Duarte, B., e Oliveira, N. M. C. (2006). Dynamic Estimation and Uncertainty Quantification for Model-Based Control of Discrete Systems. In *Proceedings of the International Symposium on Advanced Control of Chemical Processes, ADCHEM 2006*, pages 585–590, Gramado, Brazil.
- Gândara, J. F. M. e Oliveira, N. M. C. (2001). Model Predictive Control of Large–

- Scale Systems: Application to the Tennessee Eastman Process. In *Proceedings of the European Control Conference, ECC'2001*, pages 2285–2290.
- Geering, H. P. (2007). *Optimal Control with Engineering Applications*. Springer, Berlin.
- Gerds, M. (2005). Gradient evaluation in DAE optimal control problems by sensitivity equations and adjoint equations. *PAMM*, 5(1):43–46.
- Gill, P. E., Murray, W., e Saunders, M. A. (1997a). User's Guide for SNOPT 5.3: A Fortran Package for Large-Scale Nonlinear Programming. Technical Report 97-5, Department of Mathematics, University of California, San Diego.
- Gill, P. E., Murray, W., e Saunders, M. A. (1997b). User's Guide For SQOPT 5.3: a Fortran Package for Large-Scale Linear and Quadratic Programming. Technical Report NA 97-4, Department of Mathematics, University of California, San Diego.
- Gill, P. E., Murray, W., Saunders, M. A., e Wright, M. H. (1983). Computing Forward-difference Intervals for Numerical Optimization. *SIAM Journal on Scientific and Statistical Computing*, 4(2):310–312.
- Gill, P. E., Murray, W., Saunders, M. A., e Wright, M. H. (1987). Maintaining LU factors of a general sparse matrix. *Linear Algebra and its Applications*, 88-89:239–270.
- Gill, P. E., Murray, W., e Wright, M. H. (1981). *Practical Optimization*. Academic Press.
- Golshan, M., Boozarjomehry, R. B., e Pishvaie, M. R. (2005). A new approach to real time optimization of the Tennessee Eastman challenge problem. *Chemical Engineering Journal*, 112(1-3):33–44.
- Golub, G. H., Klema, V., e Stewart, G. W. (1976). Rank degeneracy and least squares problems. Technical Report STAN-CS-76-559, Computer Science Department, Stanford University.
- Golub, G. H. e Van Loan, C. F. (1996). *Matrix Computations*. John Hopkins University Press, Baltimore, 3rd edition.
- Gondzio, J. e Grothey, A. (2008). A New Unblocking Technique to Warmstart Interior Point Methods Based on Sensitivity Analysis. *SIAM Journal on Optimization*, 19(3):1184–1210.
- Grewal, M. S. e Andrews, A. P. (2001). *Kalman Filtering: Theory and Practice*. Wiley-Interscience, New York.
- Hartwich, A. e Marquardt, W. (2010). Dynamic optimization of the load change of a large-scale chemical plant by adaptive single shooting. *Computers & Chemical Engineering*, 34(11):1873–1889.
- Hartwich, A., Stockmann, K., Terboven, C., Feuerriegel, S., e Marquardt, W. (2010). Parallel sensitivity analysis for efficient large-scale dynamic optimization. *Optimization and Engineering*, 12(4):489–508.
- Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., e Woodward, C. S. (2005). SUNDIALS: Suite of Nonlinear and Differential/Algebraic



- Equation Solvers. *ACM Transactions on Mathematical Software*, 31(3):363–396.
- Hong, W., Wang, S., Li, P., Wozny, G., e Biegler, L. T. (2006). A quasi-sequential approach to large-scale dynamic optimization problems. *AIChE Journal*, 52(1):255–268.
- Jockenhövel, T., Biegler, L. T., e Wächter, A. (2003). Dynamic optimization of the Tennessee Eastman process using the OptControlCentre. *Computers & Chemical Engineering*, 27(11):1513–1531.
- Juricek, B. C., Seborg, D. E., e Larimore, W. E. (2001). Identification of the Tennessee Eastman challenge process with subspace methods. *Control Engineering Practice*, 9(12):1337–1351.
- Kameswaran, S. e Biegler, L. T. (2006). Simultaneous dynamic optimization strategies: Recent advances and challenges. *Computers & Chemical Engineering*, 30(10–12):1560–1575.
- Karypis, G. e Kumar, V. (1999). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1):359–392.
- Kavouras, A., Kevrekidis, I., Georgakis, C., e Kelley, C. (2005). Exploiting legacy dynamic simulators to accelerate steady state determination for chemical process plants. Technical report, Center for Research in Scientific Computation, North Carolina State University, Raleigh, North Carolina.
- Kiehl, M. (1999). Sensitivity Analysis of ODEs and DAEs – Theory and Implementation Guide. *Optimization Methods and Software*, 10(6):803–821.
- Kirches, C. e Leyffer, S. (2011). TACO —A Toolkit for AMPL Control Optimization. Technical report, Mathematics & Computer Science Division, Argonne National Laboratory.
- Kraus, T., Kühn, P., Wirsching, L., Bock, H. G., e Diehl, M. (2006). A Moving Horizon State Estimation algorithm applied to the Tennessee Eastman Benchmark Process. In *Proc. of IEEE Robotics and Automation Society conference on Multisensor Fusion and Integration for Intelligent Systems*.
- Kunkel, P. e Mehrmann, V. (1998). Regular Solutions of Nonlinear Differential Algebraic Equations and their Numerical Determination. *Numerical Mathematics*, 79:581–600.
- Kwon, W. e Han, S. (2005). *Receding Horizon Control - Model Predictive Control for State Models*. Advanced Textbooks in Control and Signal Processing. Springer London.
- Larsson, T., Hestetun, K., Hovland, E., e Skogestad, S. (2001). Self-Optimizing Control of a Large-Scale Plant: The Tennessee Eastman Process. *Industrial & Engineering Chemistry Research*, 40(22):4889–4901.
- Larsson, T. e Skogestad, S. (2000). Plantwide control - A review and a new design



- procedure. *Modeling, Identification and Control*, 21(4):209–240.
- Lawson, G., Aanson, R., Kincaid, D., e Krogh, F. (1979). Basic Linear Algebra Subprograms for Fortran Usage. *ACM Transactions on Mathematical Software*, 5(3):308—371.
- Leineweber, D. (1999). *Efficient Reduced SQP Methods for the Optimization of Chemical Processes Described by Large Sparse DAE Models*. PhD thesis, Universität Heidelberg.
- Leineweber, D. B., Bauer, I., Bock, H. G., e Schlöder, J. P. (2003). An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization. Part 1: theoretical aspects. *Computers & Chemical Engineering*, 27(2):157–166.
- Lewis, F. L. e Syrmos, V. L. (1995). *Optimal Control*. John Wiley & Sons, 2nd edition.
- Li, N. e Saad, Y. (2006). MIQR: A Multilevel Incomplete QR Preconditioner for Large Sparse Least-Squares Problems. *SIAM Journal on Matrix Analysis and Applications*, 28(2):524.
- Li, S. e Petzold, L. (1999). Design of new DASPK for sensitivity analysis. Technical report, University of California, Santa Barbara.
- Logsdon, J. S. e Biegler, L. T. (1989). Accurate solution of differential-algebraic optimization problems. *Industrial & Engineering Chemistry Research*, 28(11):1628–1639.
- Luyben, M. L., Tyreus, B. D., e Luyben, W. L. (1997). Plantwide control design procedure. *AIChE Journal*, 43(12):3161–3174.
- Luyben, W. L. (1996). Simple Regulatory Control of the Eastman Process. *Industrial & Engineering Chemistry Research*, 35(10):3280–3289.
- Lyman, P. R. e Georgakis, C. (1995). Plant-Wide Control of the Tennessee Eastman Problem. *Computers & Chemical Engineering*, 19(3):321–331.
- Maly, T. e Petzold, L. R. (1996). Numerical methods and software for sensitivity analysis of differential-algebraic systems. *Applied Numerical Mathematics*, 20(1-2):57–79.
- Mayne, D. Q. e Michalska, H. (1990). Receding Horizon Control of Nonlinear Systems. *IEEE Transactions on Automatic Control*, 35(7):814–824.
- Mayne, D. Q., Rawlings, J. B., V.Rao, C., e Sokaert, P. O. M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814.
- McAvoy, T. (2002). Model Predictive Statistical Process Control of Chemical Plants. *Industrial & Engineering Chemistry Research*, 41(25):6337–6344.
- McAvoy, T. e Miller, R. (1999). Incorporating Integrating Variables into Steady-State Models for Plantwide Control Analysis and Design. *Industrial & Engineering Chemistry Research*, 38(2):412–420.
- McAvoy, T. J. (1998). A Methodology for Screening Level Control Structures in Plantwide Control Systems. *Computers & Chemical Engineering*, 22(11):1543–1552.

- McAvoy, T. J. (1999). Synthesis of Plantwide Control Systems Using Optimization. *Industrial & Engineering Chemistry Research*, 38(8):2984–2994.
- McAvoy, T. J. e Ye, N. (1994). Base Control for the Tennessee Eastman Problem. *Computers & Chemical Engineering*, 18(5):383–413.
- Michalik, C., Hannemann, R., e Marquardt, W. (2009). Incremental single shooting - A robust method for the estimation of parameters in dynamical systems. *Computers & Chemical Engineering*, 33(7):1298–1305.
- Michalska, H. e Mayne, D. Q. (1993). Robust Receding Horizon Control of Constrained Nonlinear Systems. *IEEE Transactions on Automatic Control*, 38(11):1623–1633.
- Morari, M. e Lee, J. H. (1999). Model Predictive Control: Past, Present and Future. *Computers & Chemical Engineering*, 23:667–682.
- Moré, J. J. e Thuente, D. J. (1994). Line Search Algorithms With Guaranteed Sufficient Decrease. *ACM Transactions on Mathematical Software*, 20(3):286–307.
- Nocedal, J. e Wright, S. J. (1999). *Numerical Optimization*. Springer-Verlag, New York.
- Oliveira, N. e Biegler, L. (1994a). A Trust Region Strategy for Newton-type Process Control. In *IFAC Symposium on Advanced Control of Chemical Processes, ADCHEM'94*, pages 283–288, Kyoto, Japan.
- Oliveira, N. M. C. (1994). *Newton-type Algorithms for Nonlinear Constrained Chemical Process Control*. PhD thesis, Carnegie Mellon University, Pittsburgh.
- Oliveira, N. M. C. e Biegler, L. T. (1994b). Constraint handing and stability properties of model-predictive control. *AIChE Journal*, 40(7):1138–1155.
- Pantelides, C. C., Gritsis, D., Morison, K. R., e Sargent, R. W. H. (1988). The Mathematical Modelling of Transient Systems Using Differential–Algebraic Equations. *Computers & Chemical Engineering*, 12(5):449–454.
- Petzold, L., Rosen, J. B., Gill, P. E., Jay, L. O., e Park, K. (1997). Numerical Optimal Control of Parabolic PDEs Using DASOPT. In Biegler, L. T., Coleman, T. F., e Conn, A. R., editors, *Large Scale Optimization with Applications, Part II: Optimal Design and Control*, volume 93 of *IMA Volumes in Mathematics and its Applications*. Springer Verlag, Berlin.
- Petzold, L. R. (1982). A Description of DASSL: a Differential/Algebraic System Solver. Technical report, Sandia National Laboratories, Albuquerque, New Mexico.
- Price, R. M., Lyman, P. R., e Georgakis, C. (1994). Throughput Manipulation in Plantwide Control Structures. *Industrial & Engineering Chemistry Research*, 33(5):1197–1207.
- Qin, S. e Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764.

- Qin, S. J. e Badgwell, T. A. (1997). An overview of industrial model predictive control technology. In *AIChE Symposium Series*, volume 93, pages 232–256. Citeseer.
- Qin, S. J. e Badgwell, T. A. (2000). An overview of nonlinear model predictive control applications. *Nonlinear model predictive control*.
- Rao, C. V. (2000). *Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems*. PhD thesis, University of Wisconsin–Madison.
- Rao, C. V. e Rawlings, J. B. (2002). Constrained Process Monitoring: Moving–Horizon Approach. *AIChE Journal*, 48(1):97–109.
- Rao, C. V., Wright, S. J., e Rawling, J. B. (1998). Application of interior–point methods to model predictive control. *Journal of Optimization Theory and Applications*, 99:723–757.
- Rao, M. e Qiu, H. (1993). *Process Control Engineering*. Gordon an Breach Science Publishers, Amsterdam.
- Rawlings, J. B. e Mayne, D. Q. (2009). *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison, Wisconsin.
- Reich, S. (1992). Existence and Uniqueness Results for Nonlinear Differential Algebraic Equations. In *Berliner Seminar on Differential Algebraic Equations*, Humboldt University, Berlin.
- Ricker, N. L. (1993). Model Predictive Control of a Continuous, Two–Phase Reactor. *Journal of Process Control*, 3(2):109–123.
- Ricker, N. L. (1995). Optimal Steady–State Operation of the Tennessee Eastman Challenge Process. *Computers & Chemical Engineering*, 19(9):949–959.
- Ricker, N. L. (1996). Decentralized Control of the Tennessee Eastman Challenge Process. *Journal of Process Control*, 6(4):205–221.
- Ricker, N. L. e Lee, J. H. (1995a). Nonlinear Model Predictive Control of the Tennessee Eastman Challenge Process. *Computers & Chemical Engineering*, 19(9):961–981.
- Ricker, N. L. e Lee, J. H. (1995b). Nonlinear Modelling and State Estimation for the Tennessee Eastman Challenge Process. *Computers & Chemical Engineering*, 19(9):983–1005.
- Rosenthal, E. R. (2010). *GAMS - A User's Guide*. GAMS Development Corporation, Washington, DC.
- Saad, Y. (1994). SPARSKIT: a basic tool kit for sparse matrix computations. Technical report.
- Santos, L. O. (2001). *Multivariable Predictive Control of Chemical Processes*. PhD thesis, Faculdade de Ciências e Tecnologia da Universidade de Coimbra, Coimbra.
- Santos, L. O., Oliveira, N. M. C., e Biegler, L. T. (1995). Reliable and Efficient Optimization Strategies for Nonlinear Model Predictive Control. In *DYCORD+'95*, Proc.

- Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes, pages 33–38. Elsevier Science, England.
- Santos, L. O., Silva, D. C. M., e Oliveira, N. M. C. (1996). Towards a Robust Non-linear Model Predictive Control Formulation. In *Proceedings of the 2nd Portuguese Conference on Automatic Control, Controlo96*, pages 95–100, Porto.
- Sargent, R. W. H. (2000). Optimal control. *Journal of Computational and Applied Mathematics*, 124(1-2):361–371.
- Schäfer, A., Kühn, P., Diehl, M., Schlöder, J. P., e Bock, H. G. (2007). Fast reduced multiple shooting methods for nonlinear model predictive control. *Chemical Engineering and Processing: Process Intensification*, 46(11):1200–1214.
- Schlegel, M., Marquardt, W., Ehrig, R., e Nowak, U. (2004). Sensitivity analysis of linearly-implicit differential-algebraic systems by one-step extrapolation. *Applied Numerical Mathematics*, 48(1):83–102.
- Silva, D. C. M. (2005). *Controlo Predictivo Não-linear de Processos Químicos - Aplicação a Sistemas de Polimerização Descontínuos*. PhD thesis, Universidade de Coimbra.
- Sjöberg, J. (2008). *Optimal Control and Model Reduction of Nonlinear DAE Models*. PhD thesis, Linköping university, Linköping.
- Srinivasan, B., Palanki, S., e Bonvin, D. (2003). Dynamic optimization of batch processes: I. Characterization of the nominal solution. *Computers & Chemical Engineering*, 27(1):1–26.
- Srinivasan, G. R. e Arkun, Y. (1997). Control of the Tennessee Eastman Process Using Input–Output Models. *Journal of Process Control*, 7(5):387–400.
- Støren, S. e Hertzberg, T. (1995). The sequential linear quadratic programming algorithm for solving dynamic optimization problems - a review. *Computers & Chemical Engineering*, 19(Supplement 1):495–500.
- Tatjewski, P. (2007). *Advanced Control of Industrial Processes - Structures and Algorithms*. Advances in Industrial Control. Springer.
- Tian, Z. e Hoo, K. A. (2005). Multiple Model-Based Control of the Tennessee Eastman Process. *Industrial & Engineering Chemistry Research*, 44(9):3187–3202.
- Tsang, T. H., Himmelblau, D. M., e Edgar, T. F. (1975). Optimal Control via Collocation and Non–Linear Programming. *International Journal of Control*, 21(5):763–768.
- Ugray, Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., e Martí, R. (2005). *Metaheuristic Optimization via Memory and Evolution*, volume 30 of *Operations Research/Computer Science Interfaces Series*. Kluwer Academic Publishers, Boston.
- Unger, J., Kröner, A., e Marquardt, W. (1995). Structural analysis of differential-algebraic equation systems—theory and applications. *Computers & Chemical Engineering*, 19(8):867–882.
- Vanderbei, R. J. (1994). LOQO: an interior point code for quadratic programming.

- Technical Report SQR-94-15, Princeton University.
- Vanderbei, R. J. e Shanno, D. F. (1999). An Interior-Point Algorithm for Nonconvex Nonlinear Programming. *Computational Optimization and Applications*, 13(1–3):231–252.
- Vasantharajan, S. e Biegler, L. (1990). Simultaneous strategies for optimization of differential-algebraic systems with enforcement of error criteria. *Computers & Chemical Engineering*, 14(10):1083–1100.
- Vassiliadis, V. S., Sargent, R. W. H., e Pantelides, C. C. (1994). Solution of a Class of Multistage Dynamic Optimization Problems. 2. Problems with Path Constraints. *Industrial & Engineering Chemistry Research*, 33(9):2123–2133.
- Vinson, D., Georgakis, C., e Fossy, J. (1995). Studies in plant-wide controllability using the Tennessee Eastman Challenge problem, the case for multivariable control. In *Proceedings of the American Control Conference*, pages 250–254.
- Wächter, A. (2002). *An Interior Point Algorithm for Large-Scale Nonlinear Optimization with Applications in Process Engineering*. PhD thesis, Carnegie Mellon University.
- Wächter, A. e Biegler, L. T. (2006). On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming*, 106(1):25–57.
- Whaley, R. C., Petitet, A., e Dongarra, J. J. (2000). Automated Empirical Optimization of Software and the ATLAS Project. Technical report, University of Tennessee, Knoxville.
- Willems, J. L. (1970). *Stability Theory of Dynamical Systems*. Nelson, London.
- Wolfram, S. (1988). *Mathematica. A Systems for Doing Mathematics by Computer*. Addison-Wesley Publishing Company, Inc., Redwood City, CA.
- Zerkaoui, S., Druaux, F., Leclercq, E., e Lefebvre, D. (2010). Indirect neural control for plant-wide systems: Application to the Tennessee Eastman Challenge Process. *Computers & Chemical Engineering*, 34(2):232–243.
- Zheng, A. (1998). Nonlinear model predictive control of the Tennessee Eastman process. In *Proceedings of the 1998 American Control Conference*, pages 1700–1704.



## Apêndice A

# Comportamento das variáveis manipuláveis do processo TE

São aqui apresentados os resultados obtidos para o comportamento das variáveis manipuláveis do processo TE no interior das regiões de operação viáveis, apresentadas na Secção 3.4.3. Para as três razões entre os caudais mássicos de  $G$  e  $H$  no produto, são apresentados os valores das variáveis manipuláveis em todos os pontos considerados na fronteira e no interior das regiões. São também apresentadas as amplitudes de variação dessas variáveis, ou seja, as diferenças entre os valores máximo e mínimo de uma dada variável que correspondem aos mesmos custo e caudal mássico de  $G$  no produto.

### A.1 Regime de operação $10G/90H$

Na Figura A.1 é apresentada a variação das variáveis manipuláveis na região viável para o regime de operação  $10G/90H$ .



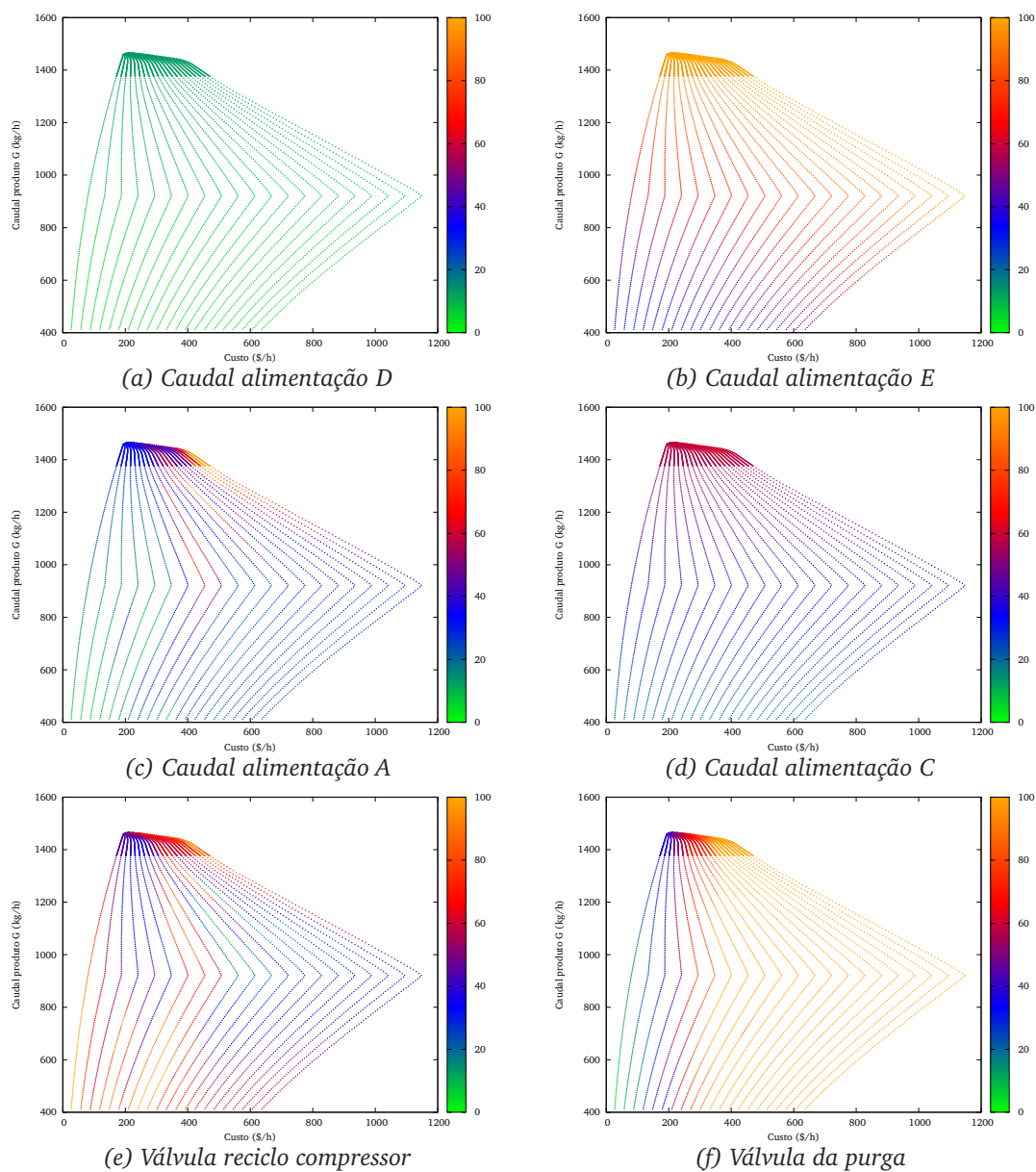


Figura A.1 Valores das variáveis manipuláveis no interior da região viável para o regime de operação 10G/90H.



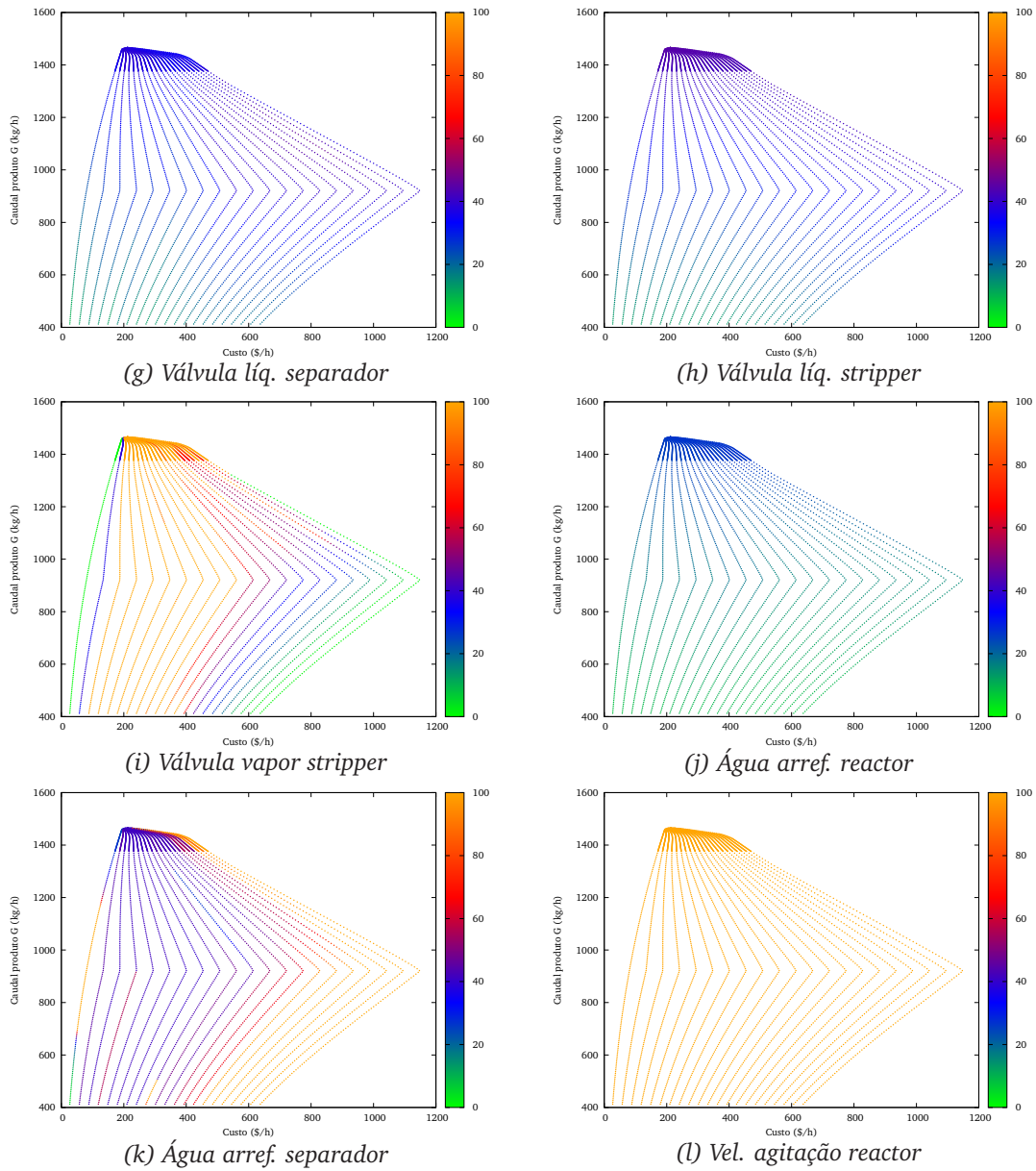
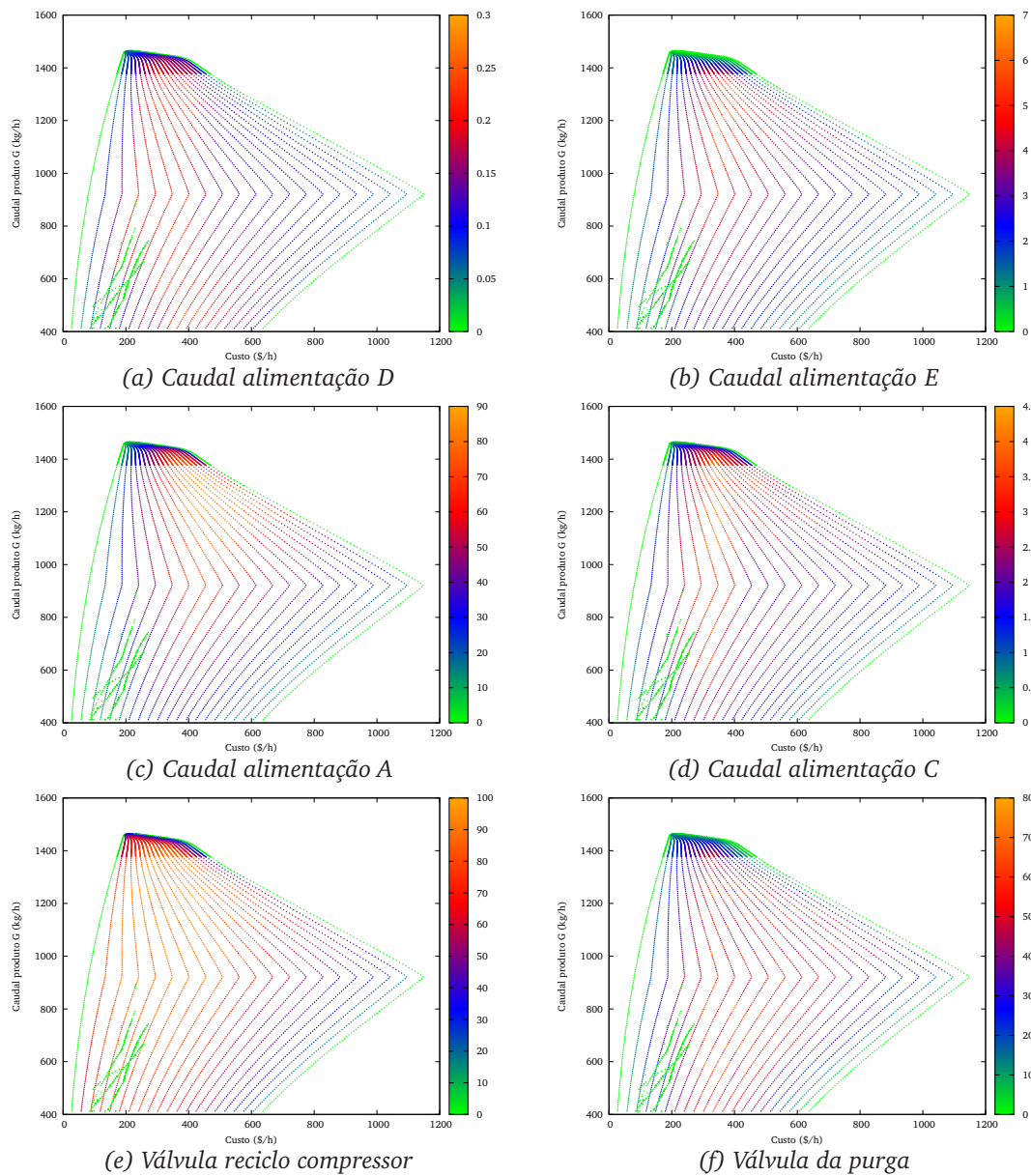


Figura A.1 (cont.) Valores das variáveis manipuláveis no interior da região viável para o regime de operação 10G/90H.

Na Figura A.2 é apresentada, para cada variável manipulável, a amplitude de variação possível, correspondente ao mesmo custo e caudal de produto G.



**Figura A.2** Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 10G/90H.

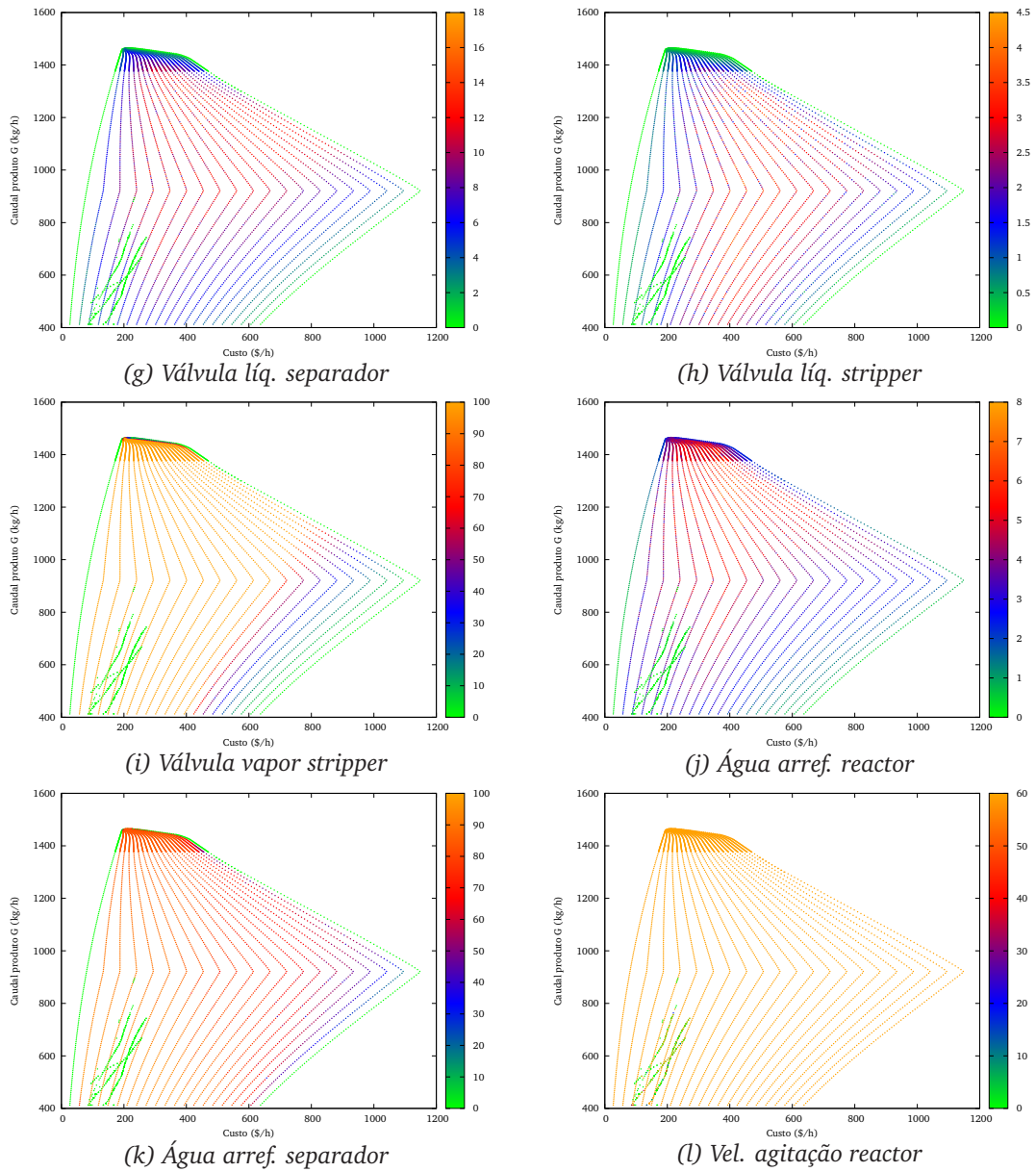
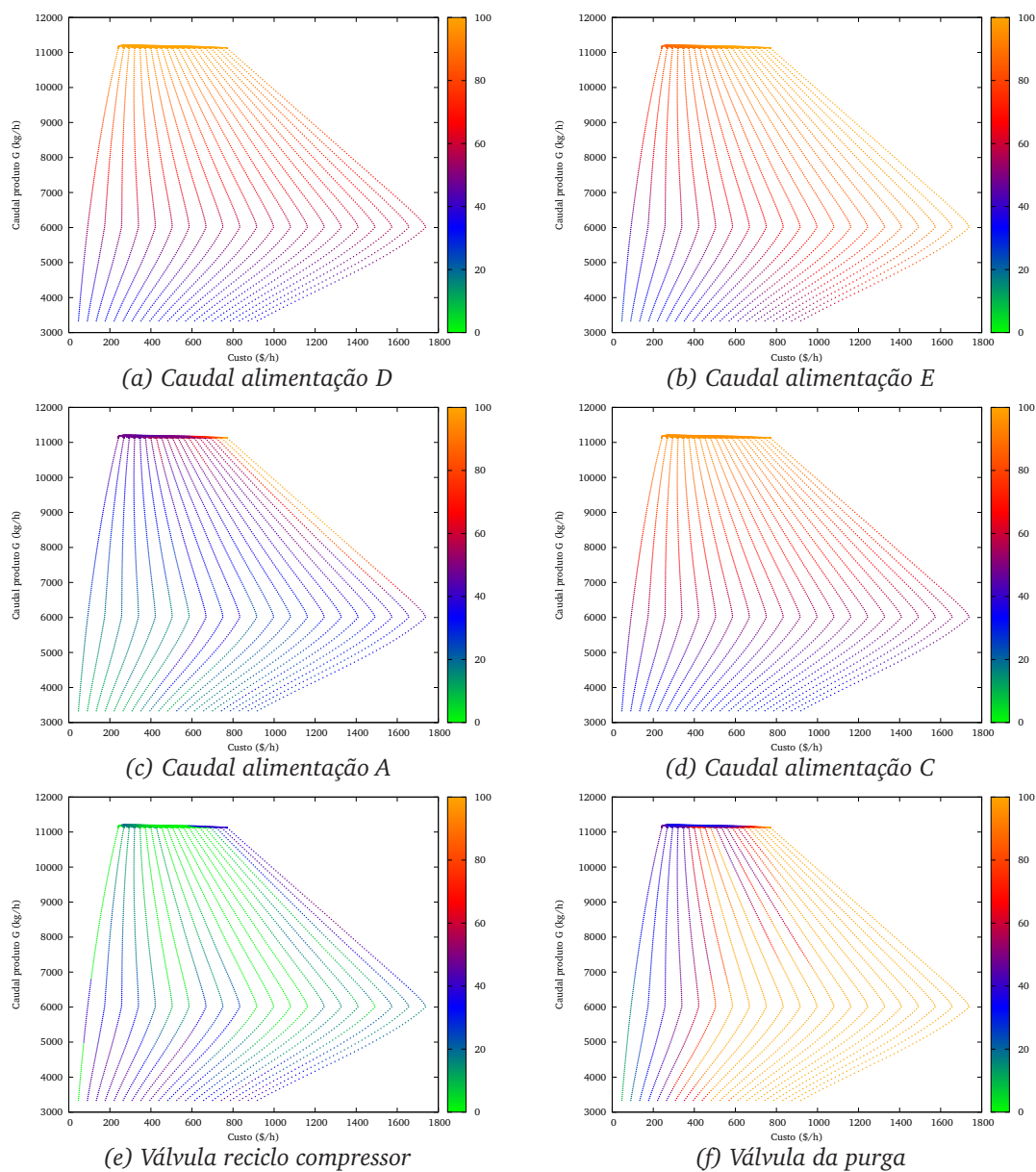


Figura A.2 (cont.) Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 10G/90H.

## A.2 Regime de operação 50G/50H

Na Figura A.3 é apresentada a variação das variáveis manipuláveis na região viável para o regime de operação 50G/50H.



**Figura A.3** Valores das variáveis manipuláveis no interior da região viável para o regime de operação 50G/50H.

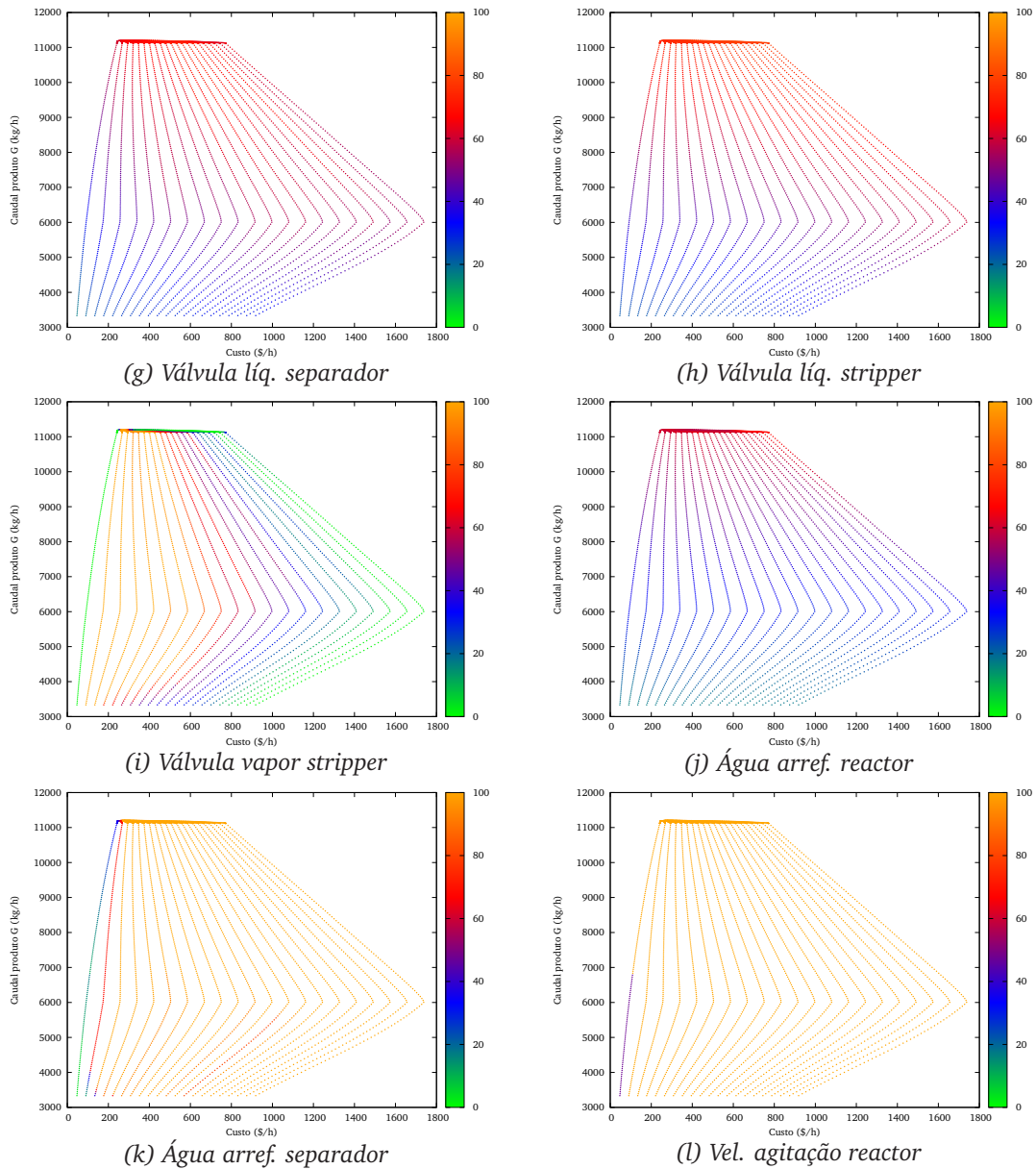
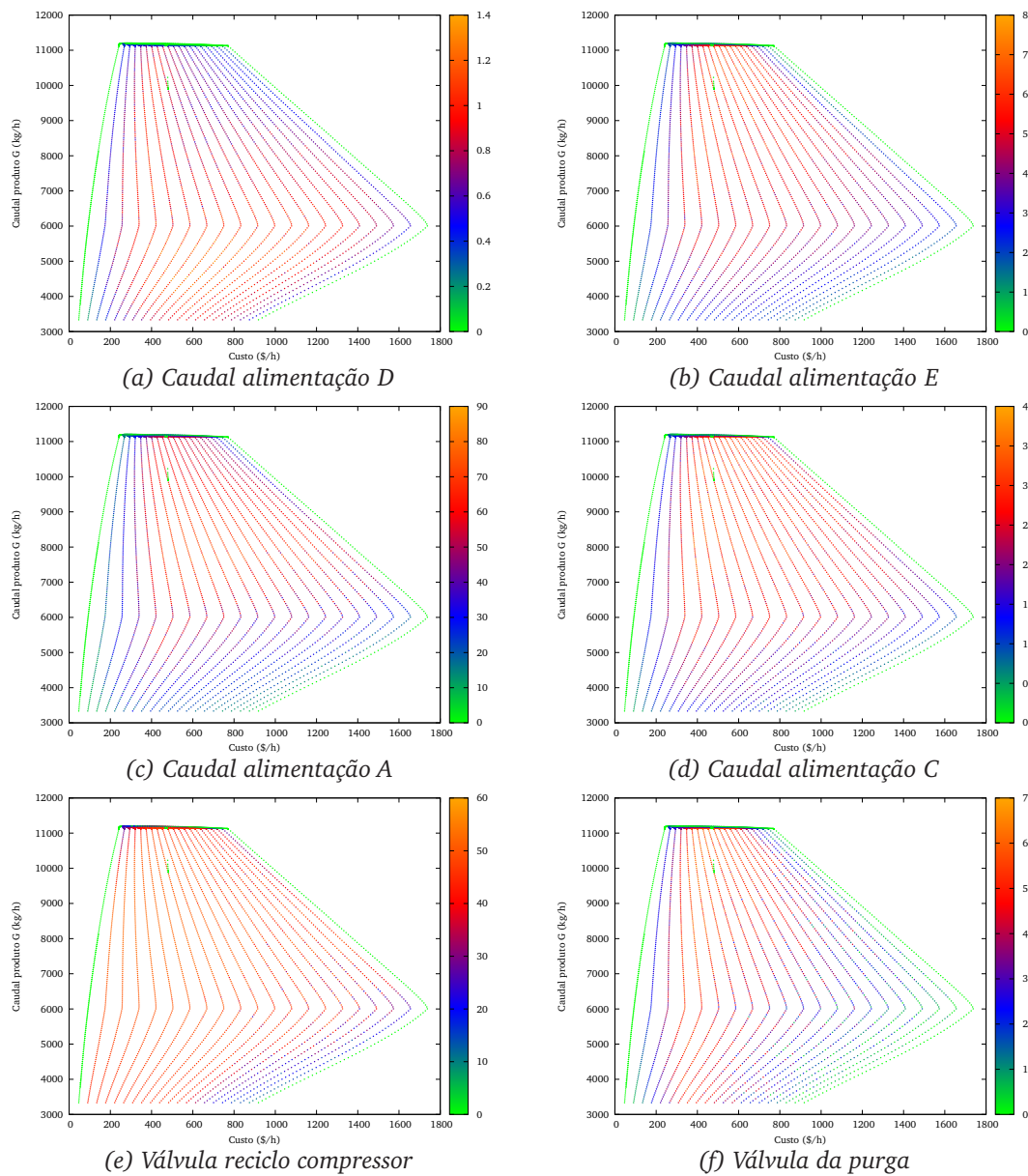


Figura A.3 (cont.) Valores das variáveis manipuláveis no interior da região viável para o regime de operação 50G/50H.

Na Figura A.4 é apresentada, para cada variável manipulável, a amplitude de variação possível, correspondente ao mesmo custo e caudal de produto G.



**Figura A.4** Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 50G/50H.

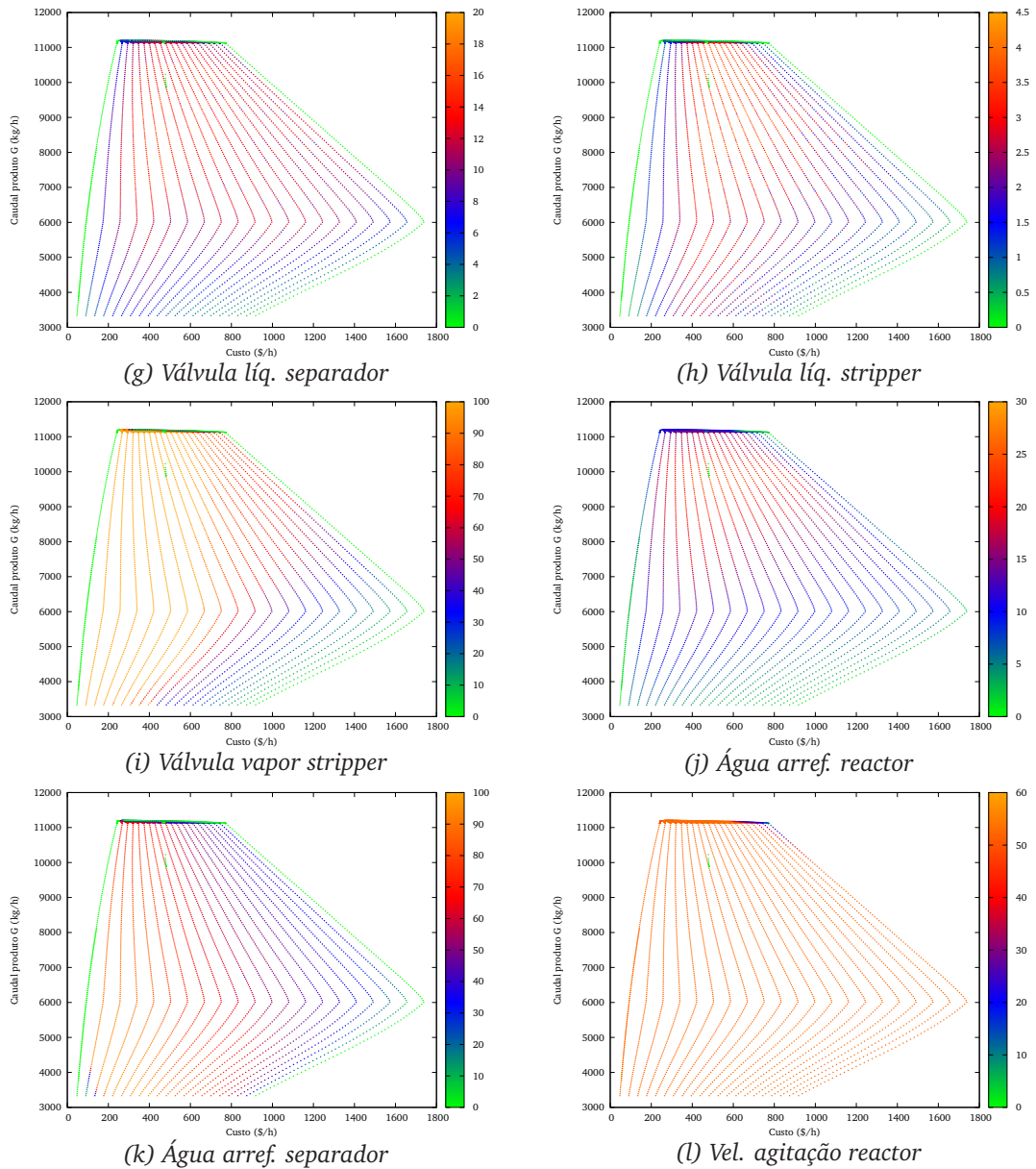
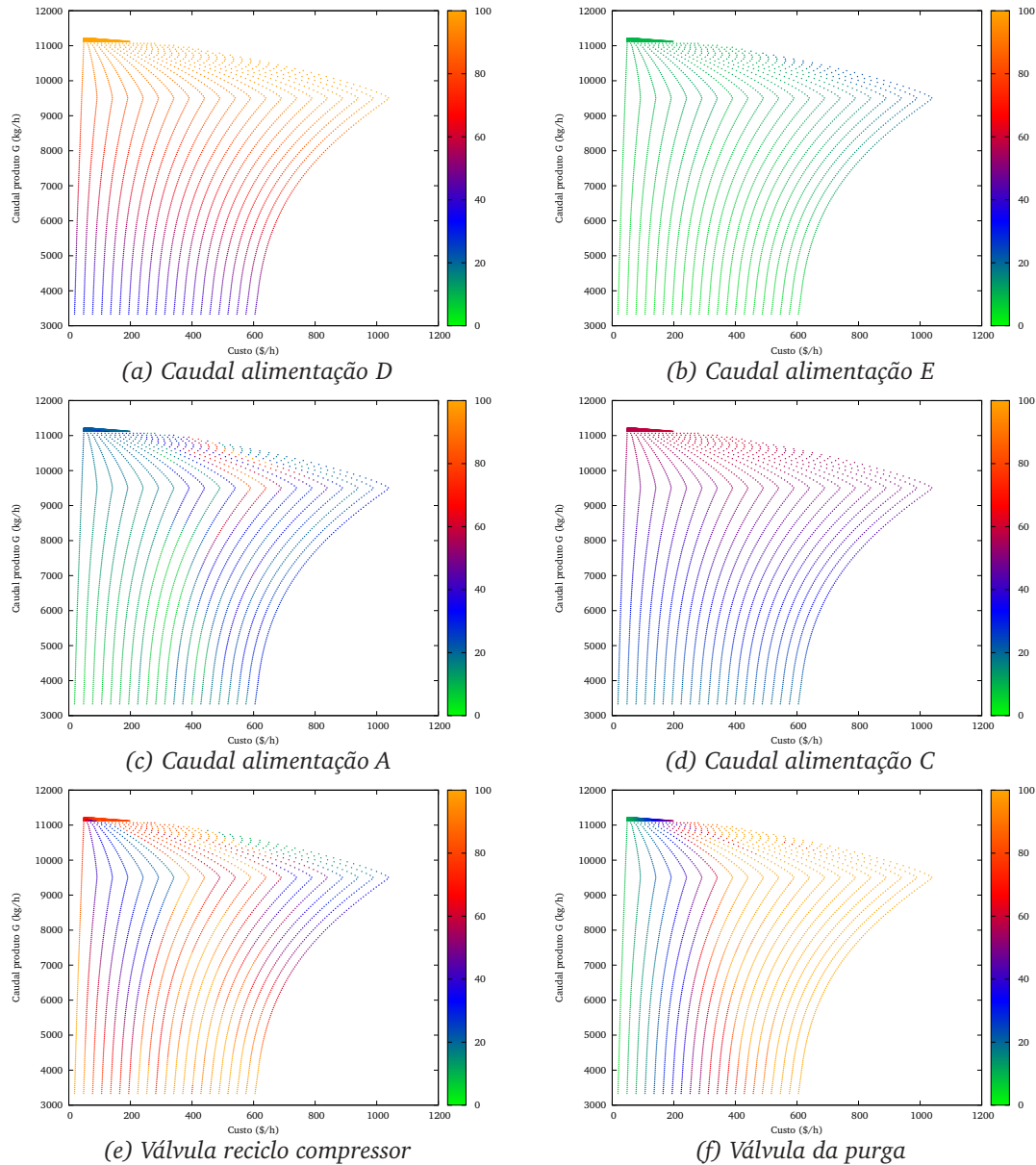


Figura A.4 (cont.) Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 50G/50H.



### A.3 Regime de operação 90G/10H

Na Figura A.5 é apresentada a variação das variáveis manipuláveis na região viável para o regime de operação 90G/10H.



**Figura A.5** Valores das variáveis manipuláveis no interior da região viável para o regime de operação 90G/10H.



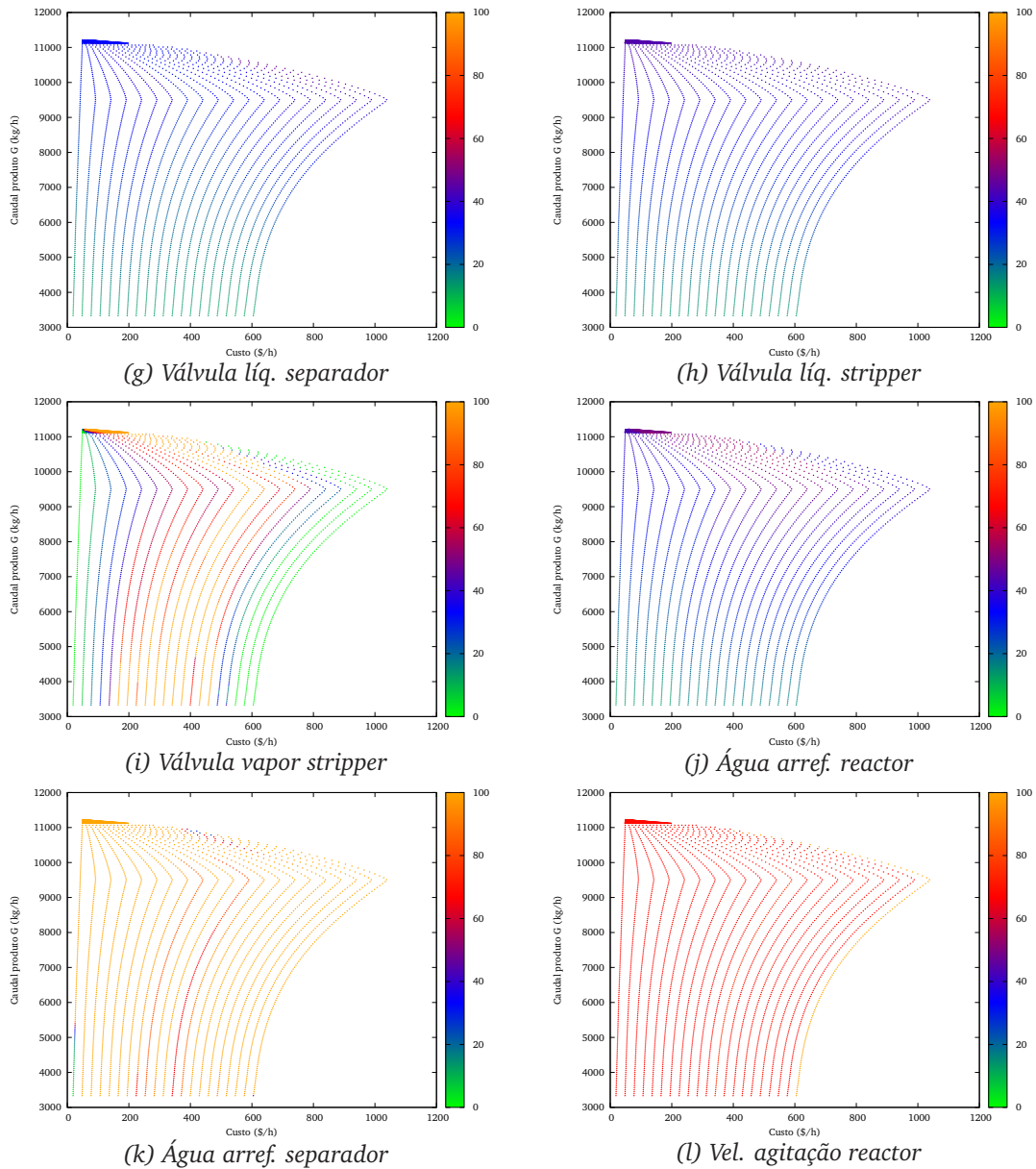
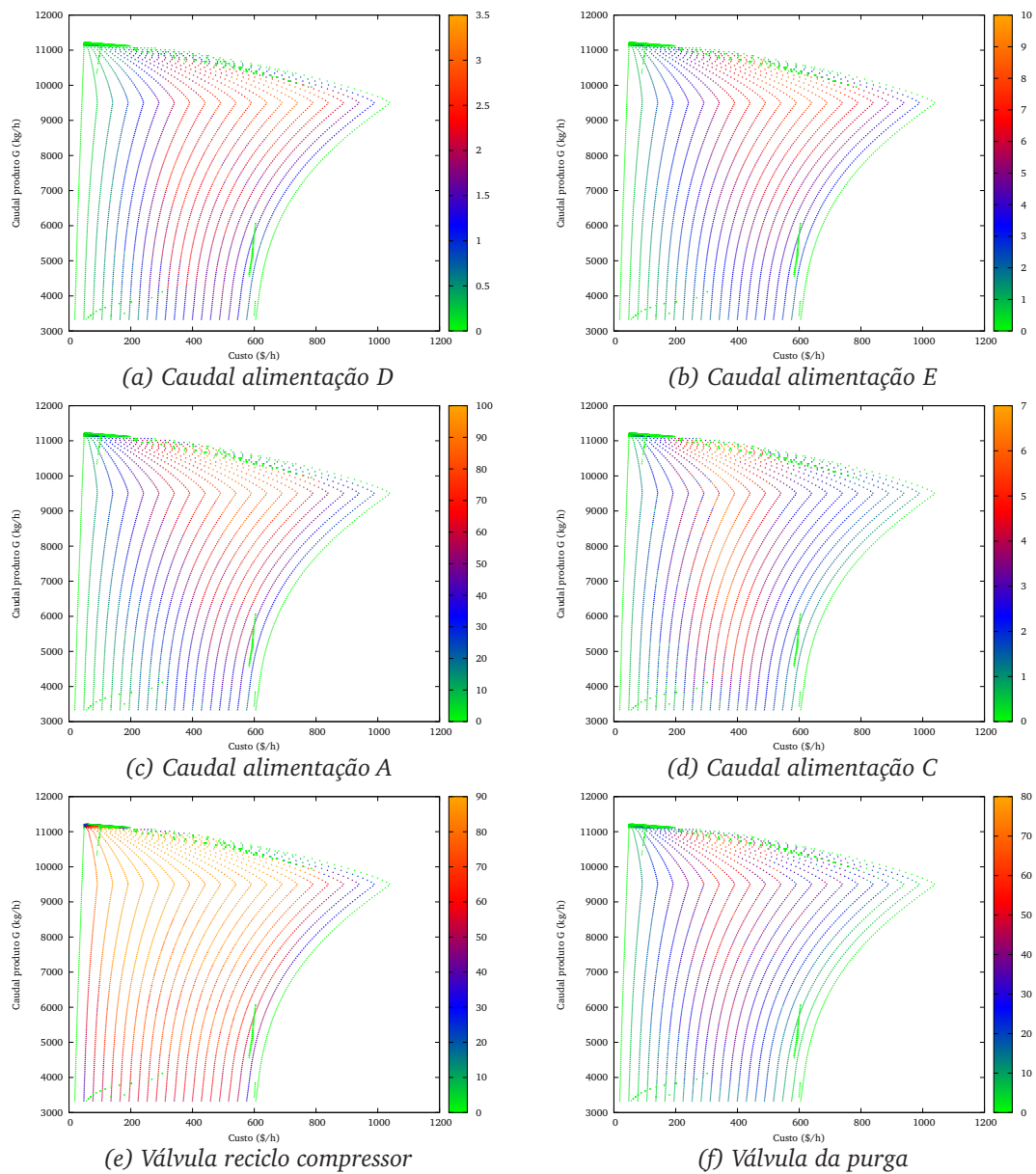


Figura A.5 (cont.) Valores das variáveis manipuláveis no interior da região viável para o regime de operação 90G/10H.

Na Figura A.6 é apresentada, para cada variável manipulável, a amplitude de variação possível, correspondente ao mesmo custo e caudal de produto G.



**Figura A.6** Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 90G/10H.

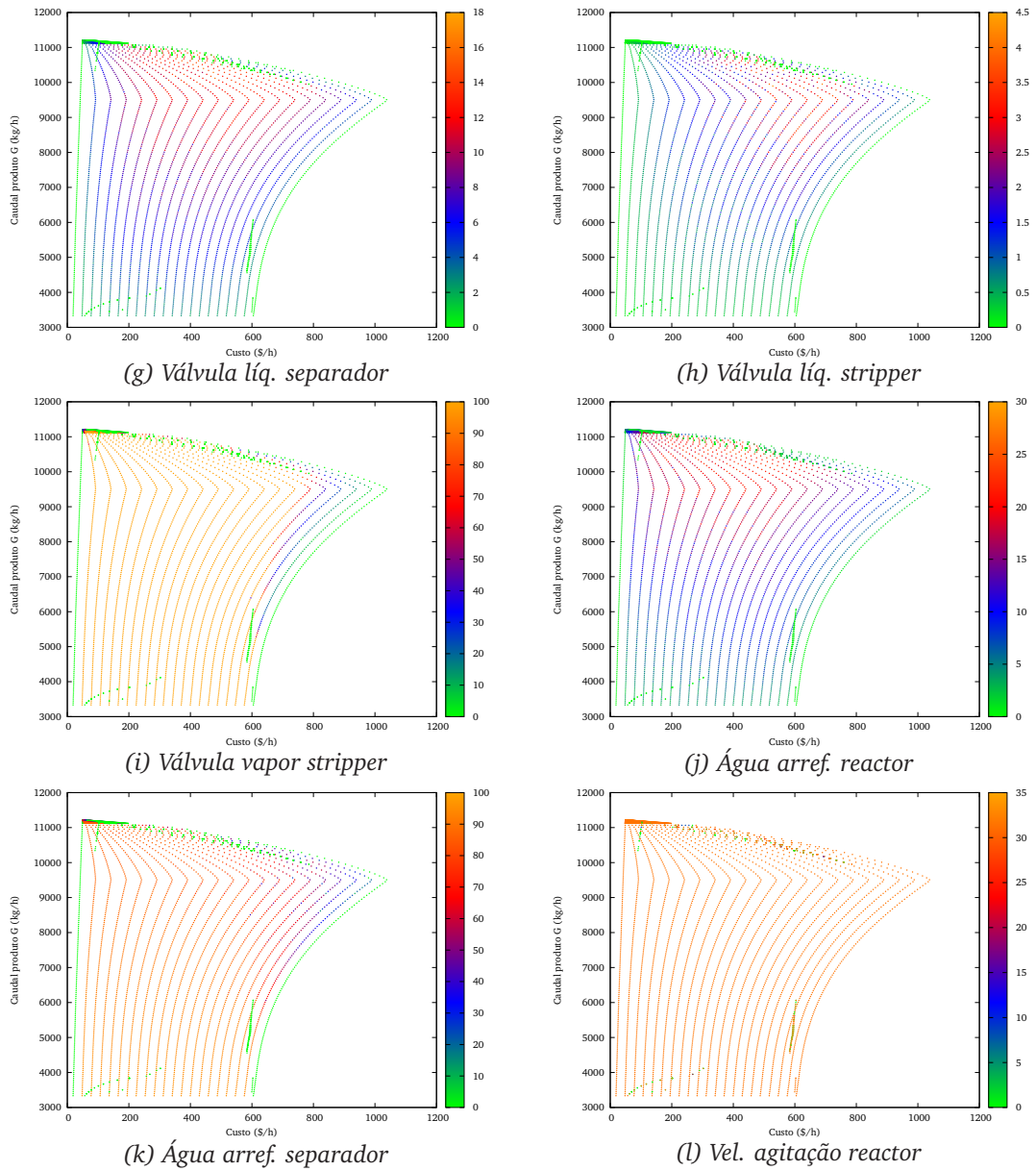


Figura A.6 (cont.) Amplitude das variáveis manipuláveis no interior da região viável para o regime de operação 90G/10H.

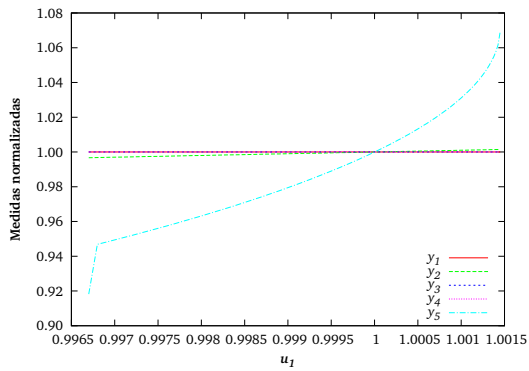


## Apêndice B

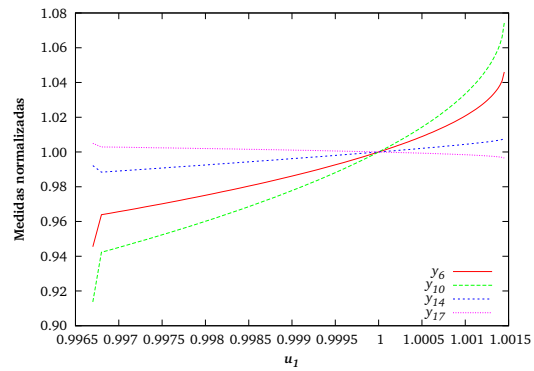
### Ganhos estacionários do processo TE

São aqui apresentados os gráficos relativos aos ganhos estacionários das variáveis medidas relativamente às variáveis manipuláveis do processo TE. Estes gráficos foram obtidos da forma descrita na Secção 3.4.4.

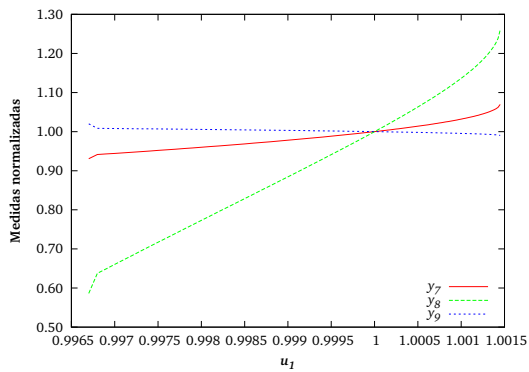
Na Figura B.1 são apresentados os resultados obtidos para o caudal de alimentação de  $D$ ,  $u_1$ . Nestes resultados, e nos seguintes, todas as variáveis são normalizadas utilizando os valores correspondentes ao modo base de operação (modo 1).



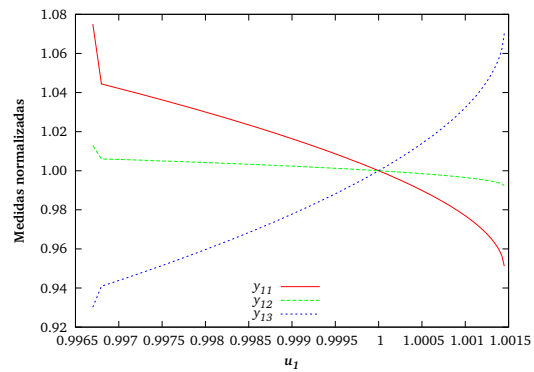
(a)



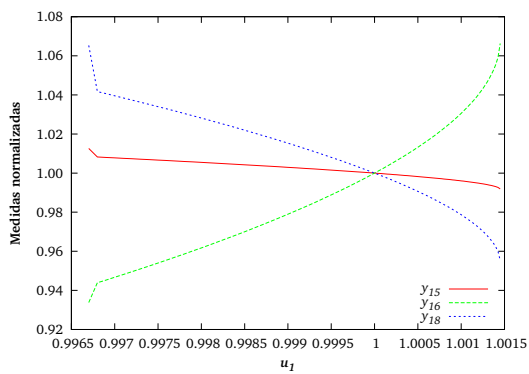
(b)



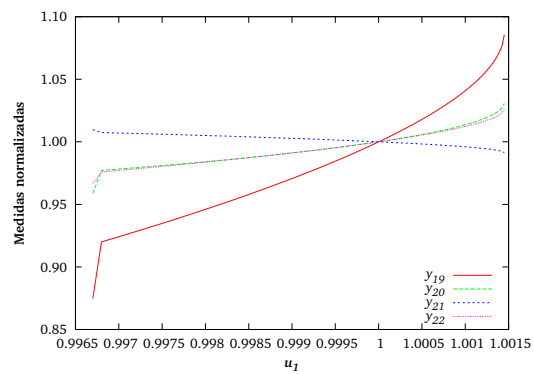
(c)



(d)

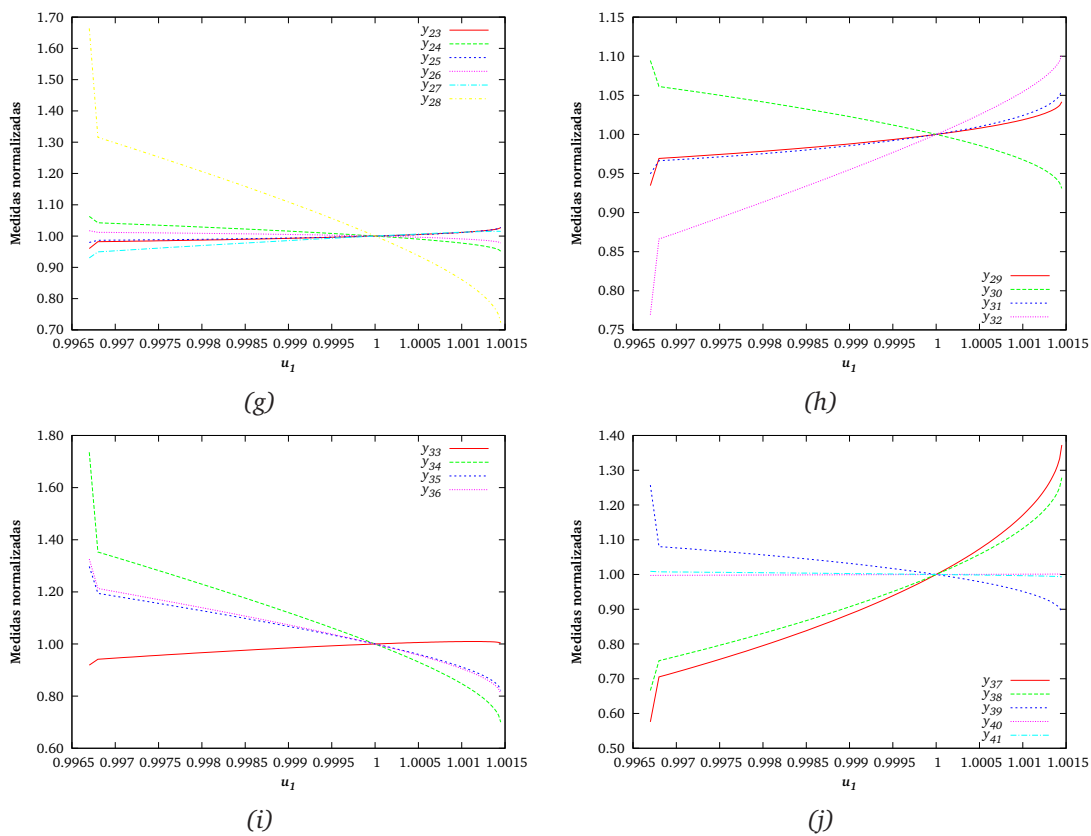


(e)



(f)

Figura B.1 Variáveis medidas em função do caudal de alimentação de  $D$  ( $u_1$ ).



**Figura B.1 (cont.)** Variáveis medidas em função do caudal de alimentação de  $D$  ( $u_1$ ).

Na Figura B.2 são apresentados os resultados para o caudal de alimentação de  $E$  ( $u_2$ ).

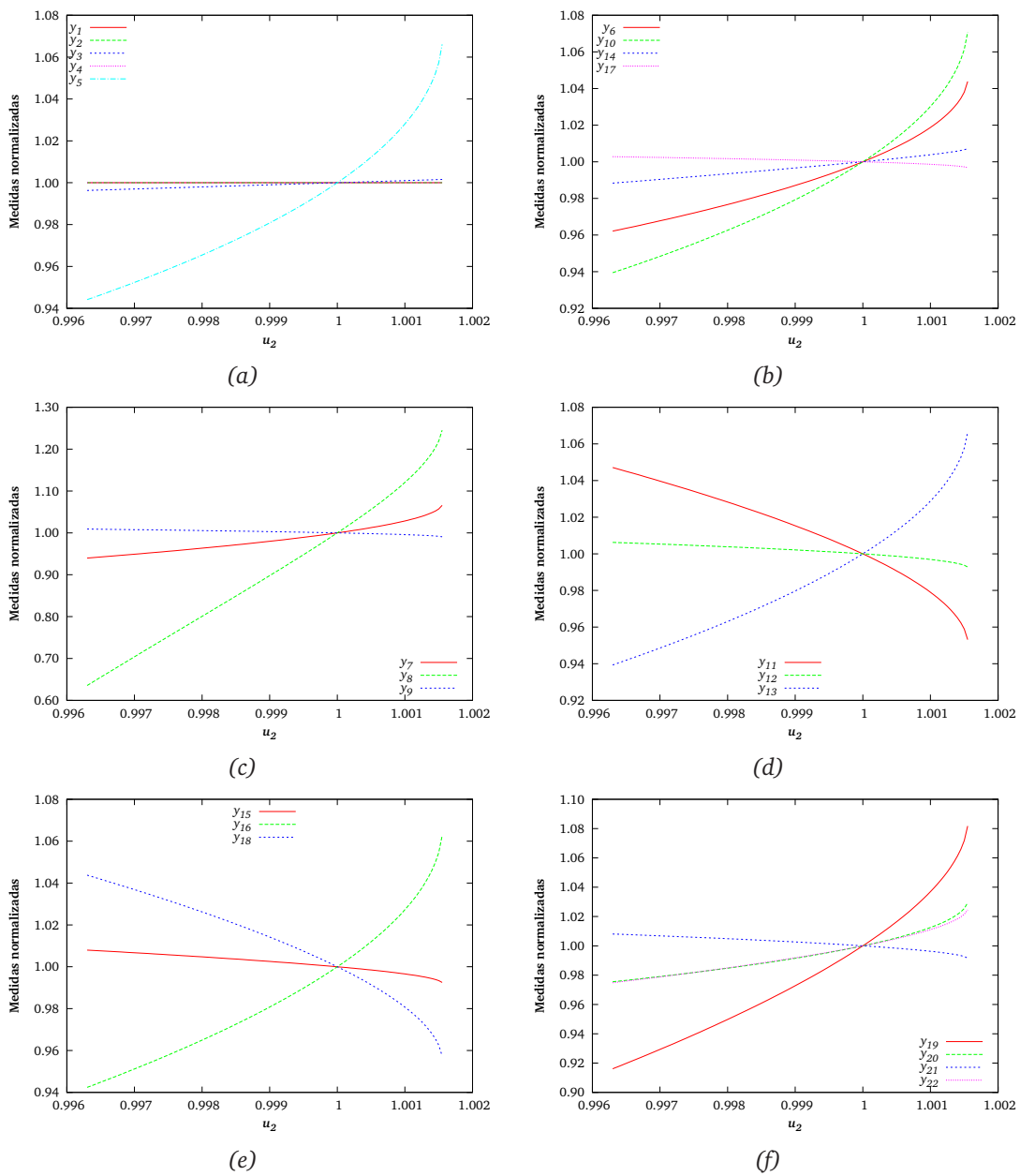
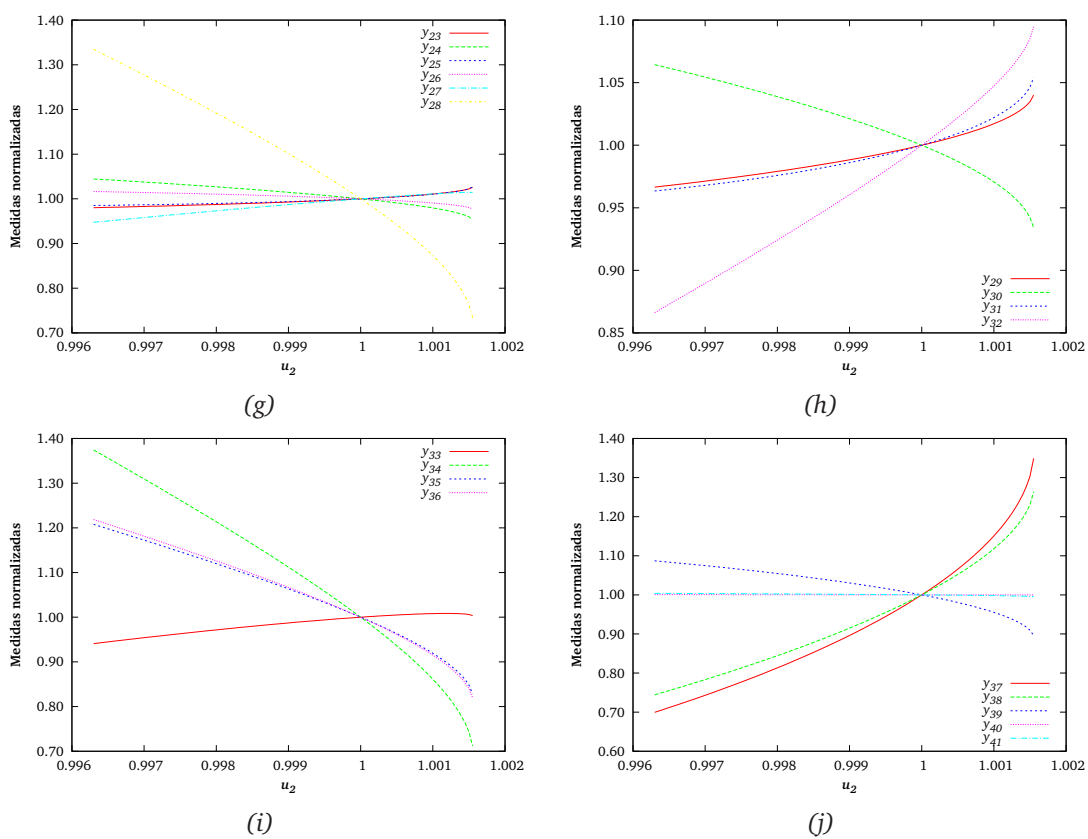


Figura B.2 Variáveis medidas em função do caudal de alimentação de  $E$  ( $u_2$ ).





**Figura B.2 (cont.)** Variáveis medidas em função do caudal de alimentação de  $E$  ( $u_2$ ).

Na Figura B.3 são apresentados os resultados relativos ao caudal de alimentação de A.

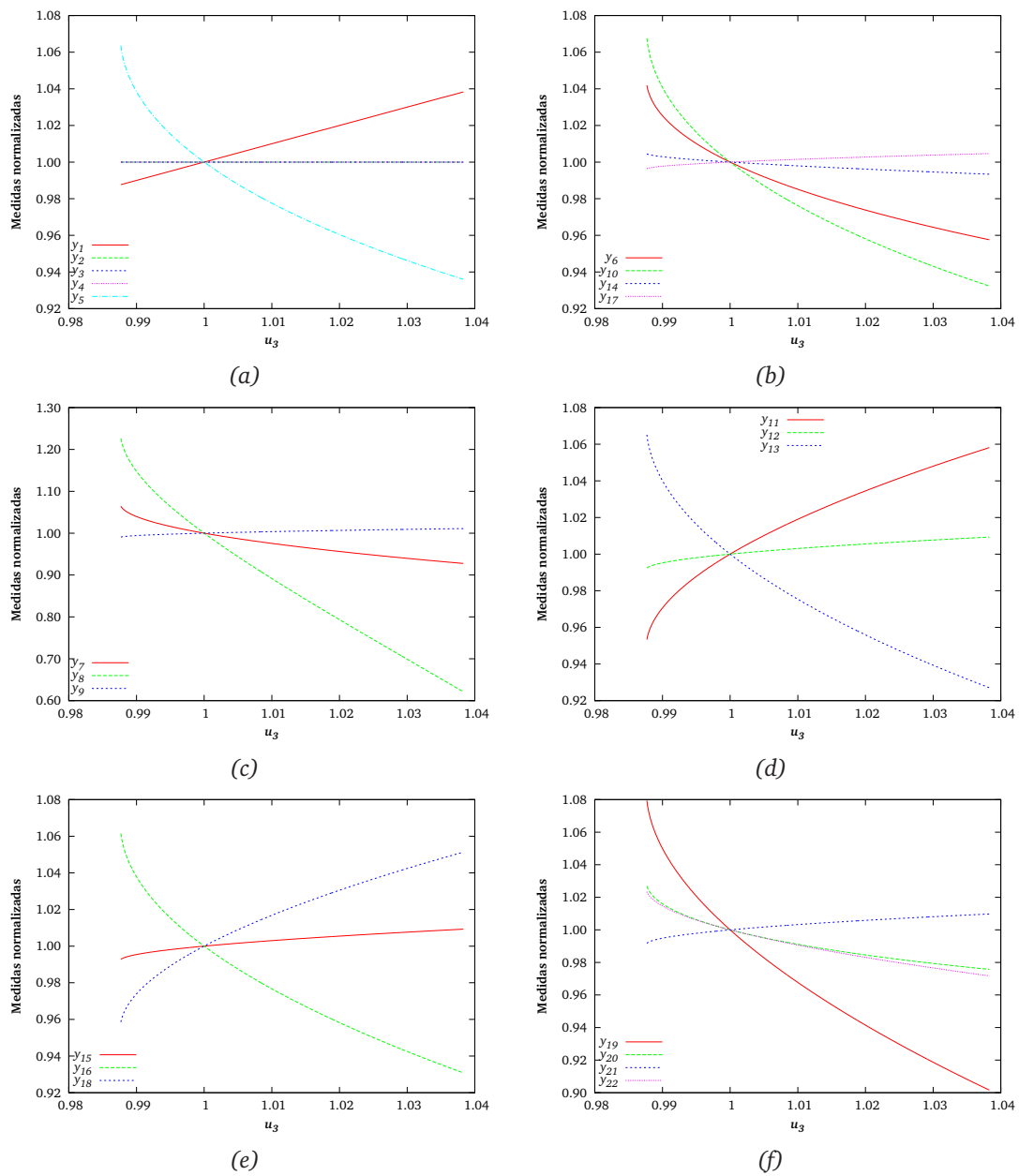


Figura B.3 Variáveis medidas em função do caudal de alimentação de A ( $u_3$ ).

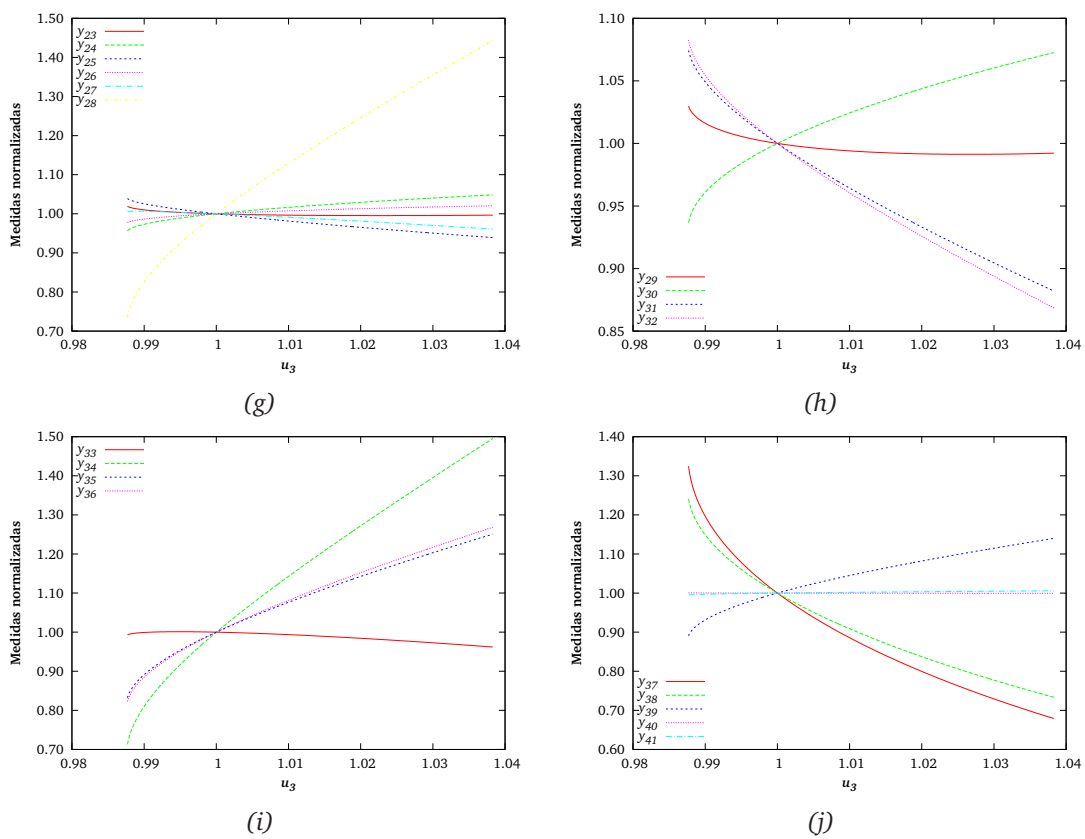


Figura B.3 (cont.) Variáveis medidas em função do caudal de alimentação de A ( $u_3$ ).

Nas Figuras B.4, B.5 e B.6 são, respectivamente, apresentados os resultados relativos ao caudal de alimentação de A e C ( $u_4$ ), à posição da válvula de reciclo do compressor ( $u_5$ ) e da posição da válvula da purga ( $u_6$ ).

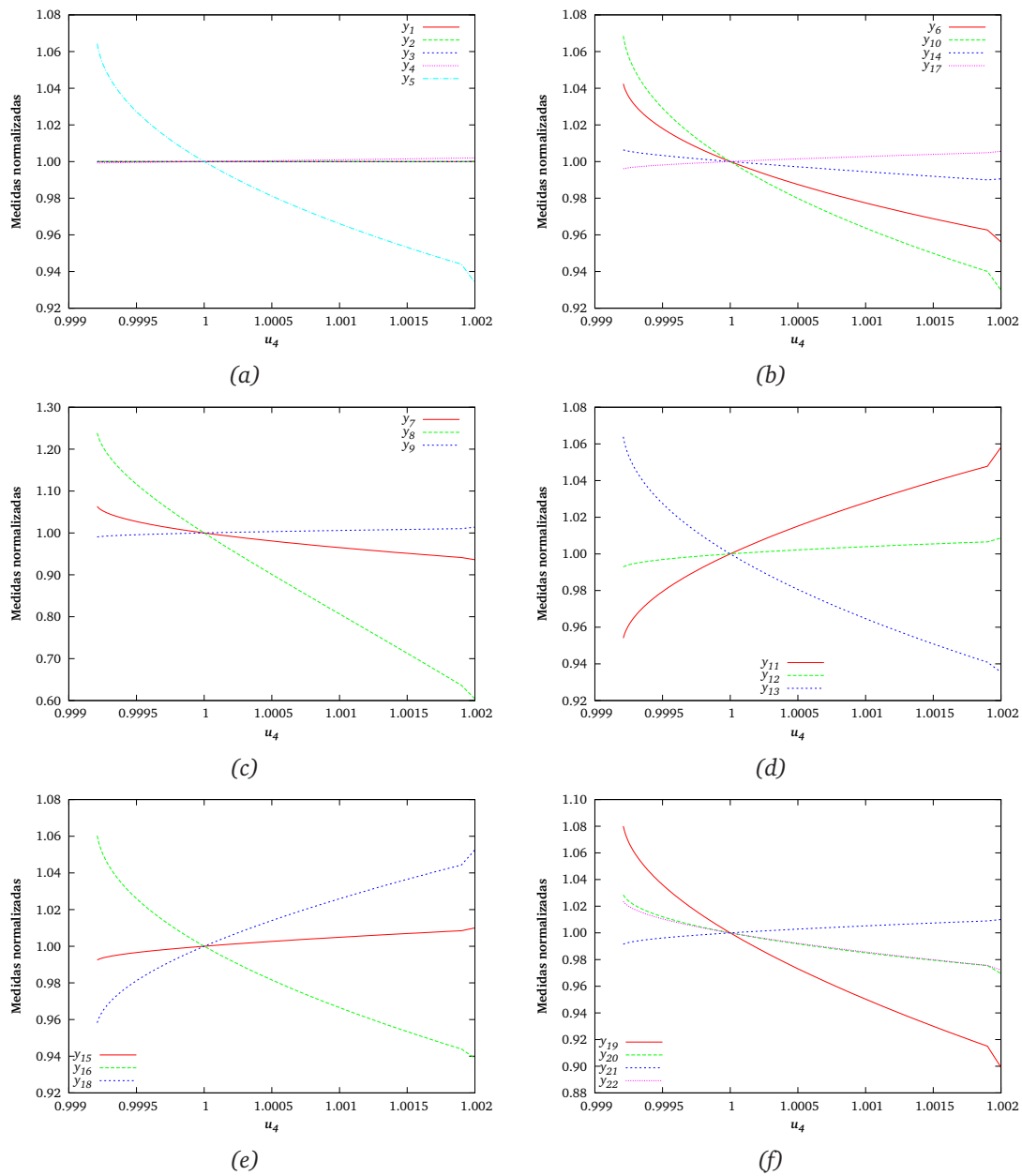
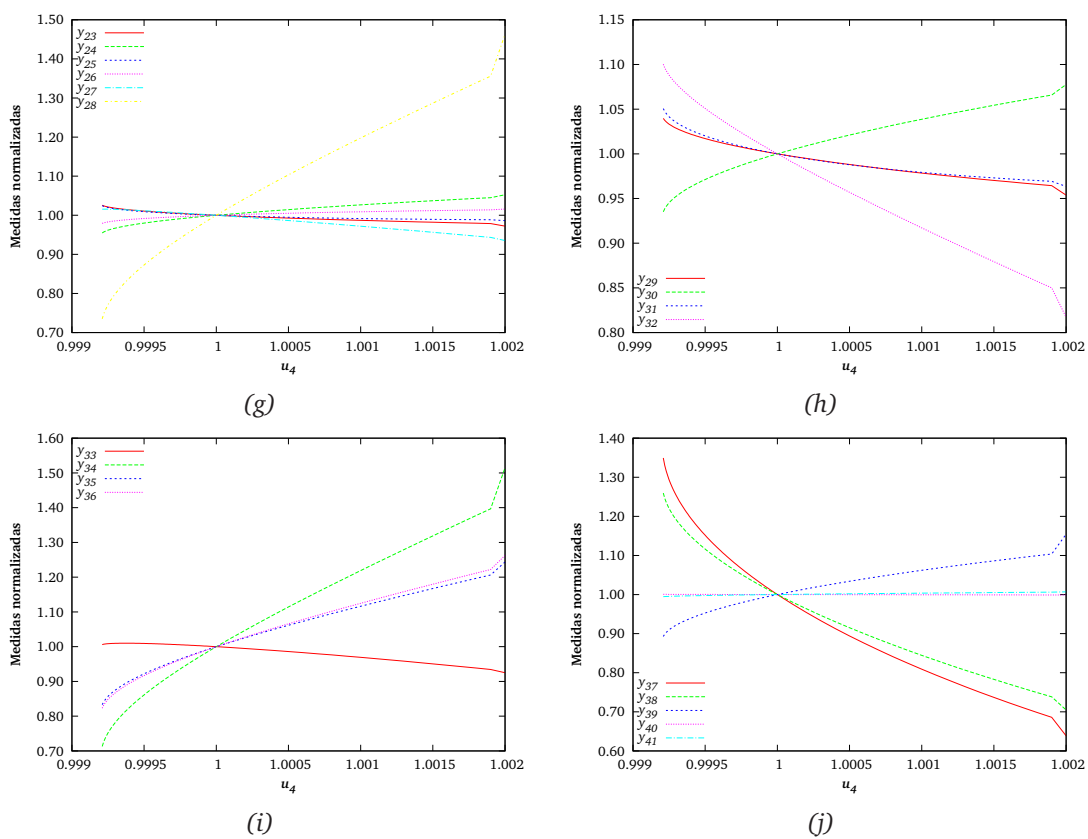


Figura B.4 Variáveis medidas em função do caudal de alimentação de A e C ( $u_4$ ).



**Figura B.4 (cont.)** Variáveis medidas em função do caudal de alimentação de A e C ( $u_4$ ).

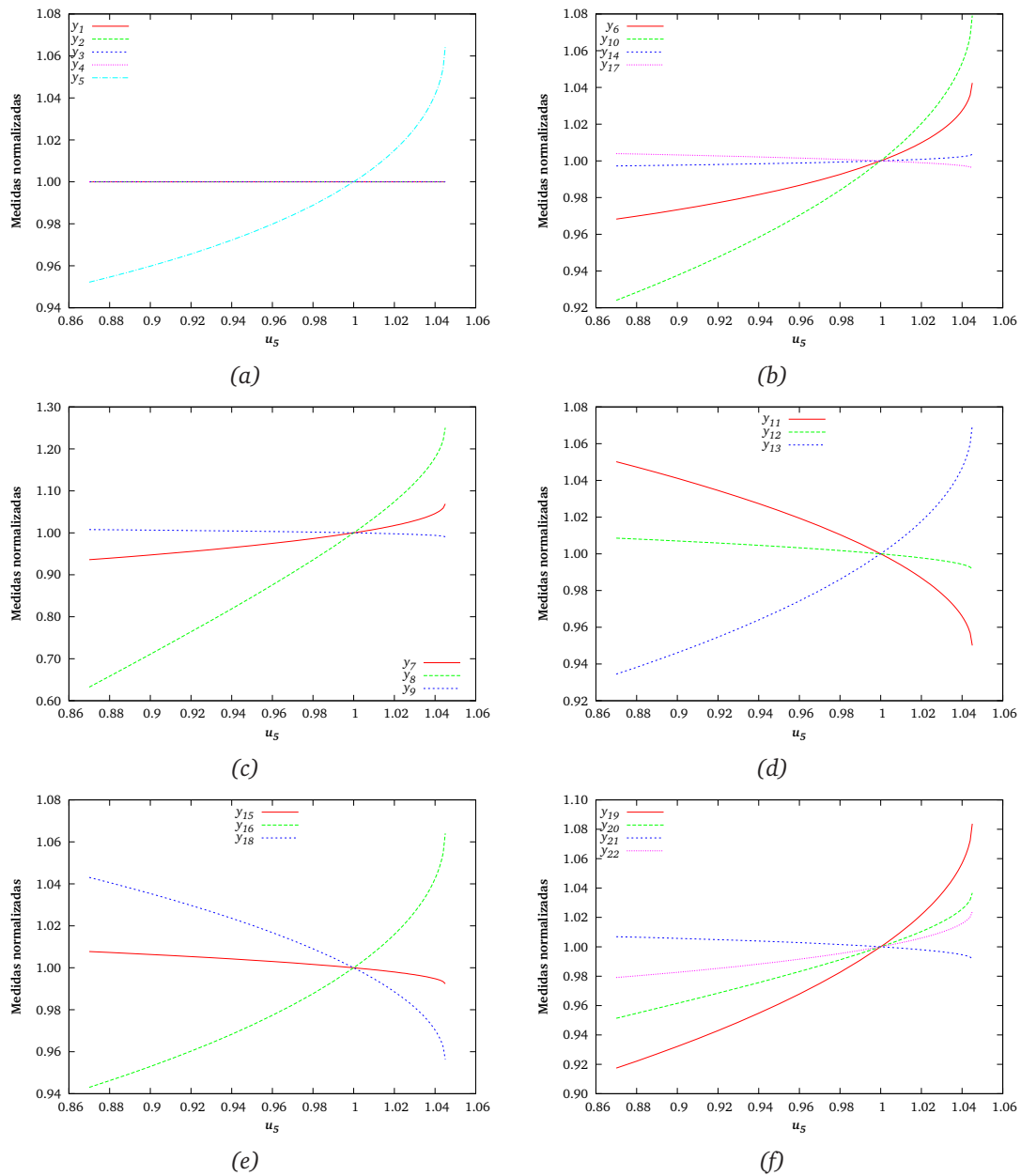
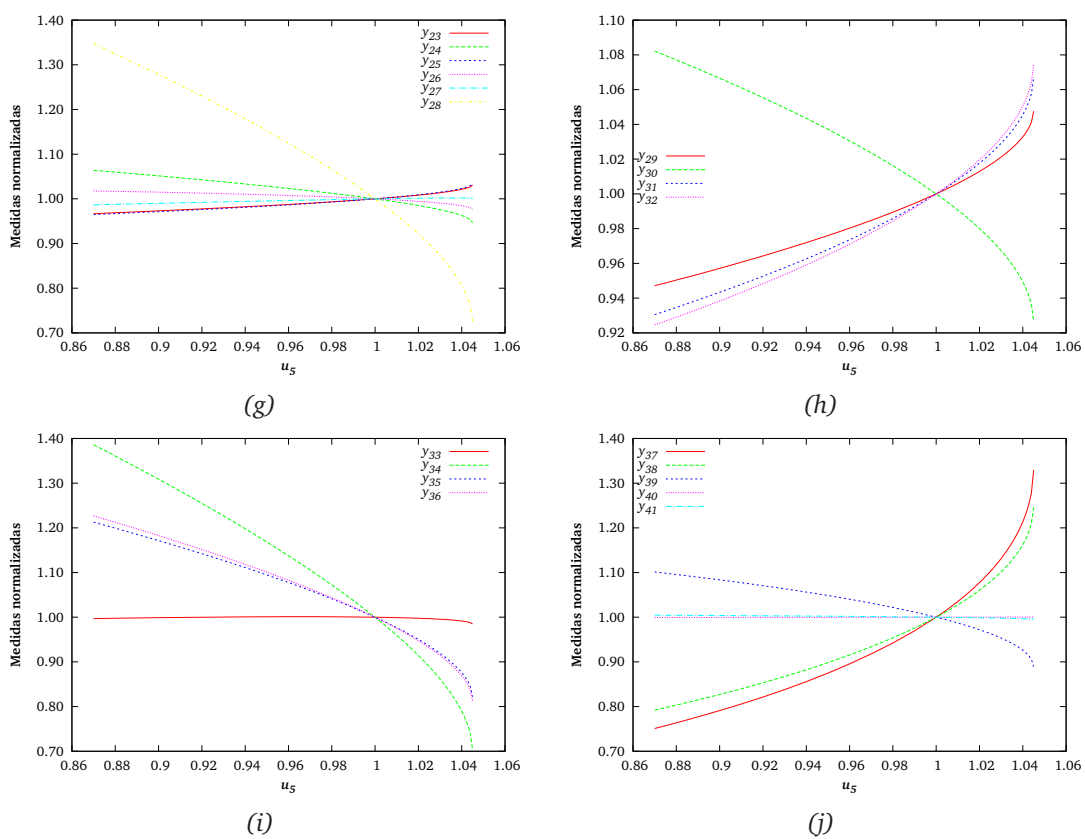
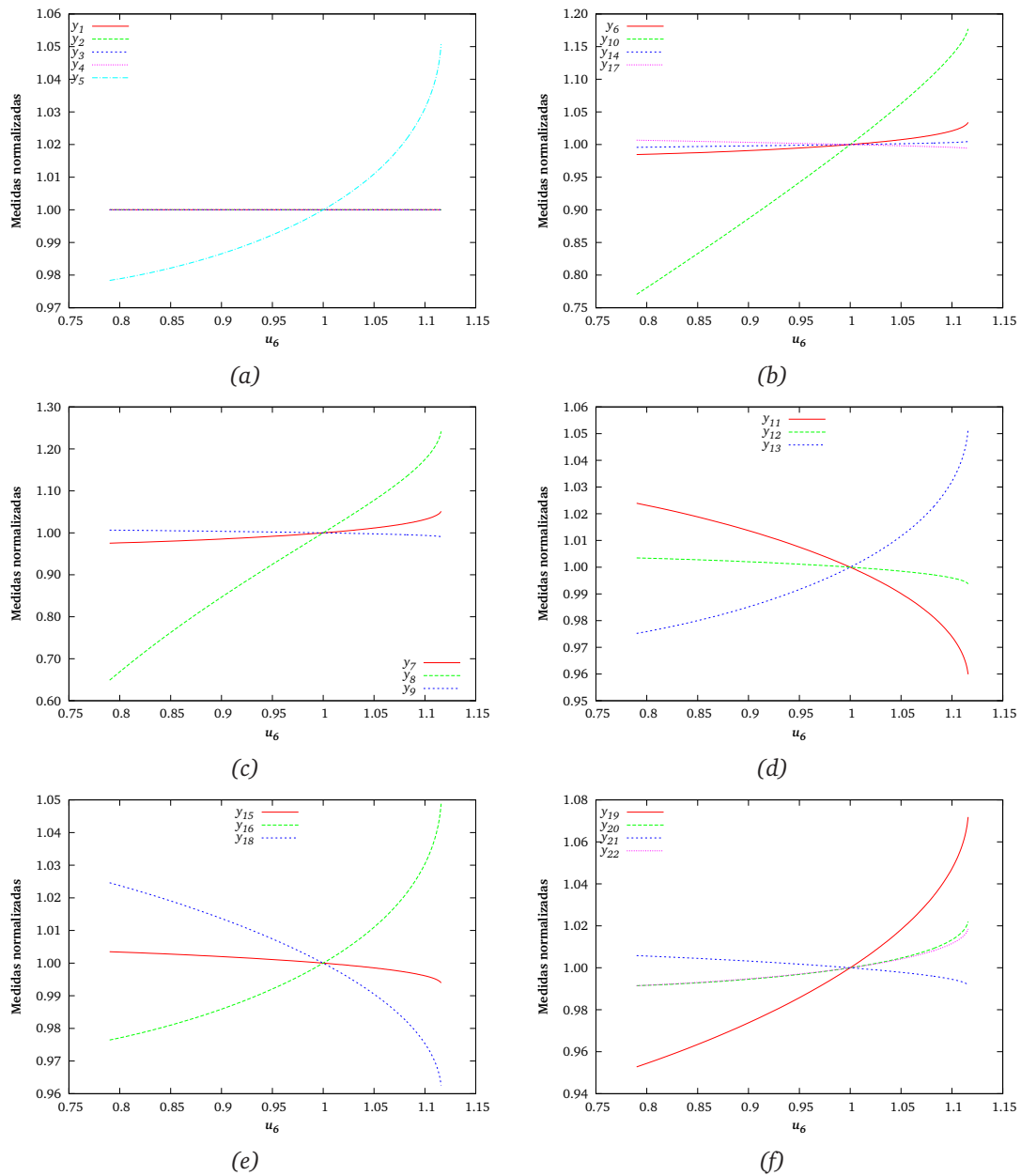


Figura B.5 Variáveis medidas em função da posição da válvula de reciclo do compressor ( $u_5$ ).

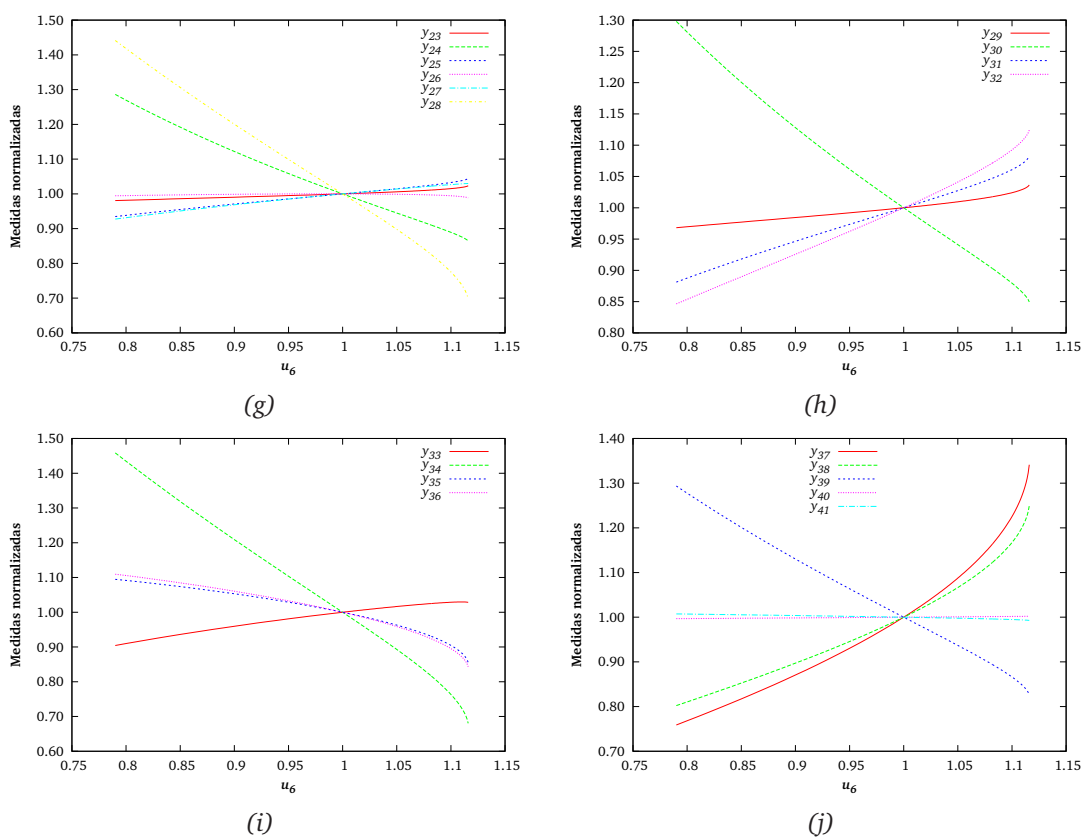


**Figura B.5 (cont.)** Variáveis medidas em função da posição da válvula de reciclo do compressor ( $u_5$ ).



**Figura B.6** Variáveis medidas em função da posição da válvula da corrente de purga ( $u_6$ ).





**Figura B.6 (cont.)** Variáveis medidas em função da posição da válvula da corrente de purga ( $u_6$ ).

Na Figura B.7 são apresentados os resultados obtidos para o caudal de vapor no *stripper*,  $u_9$ .

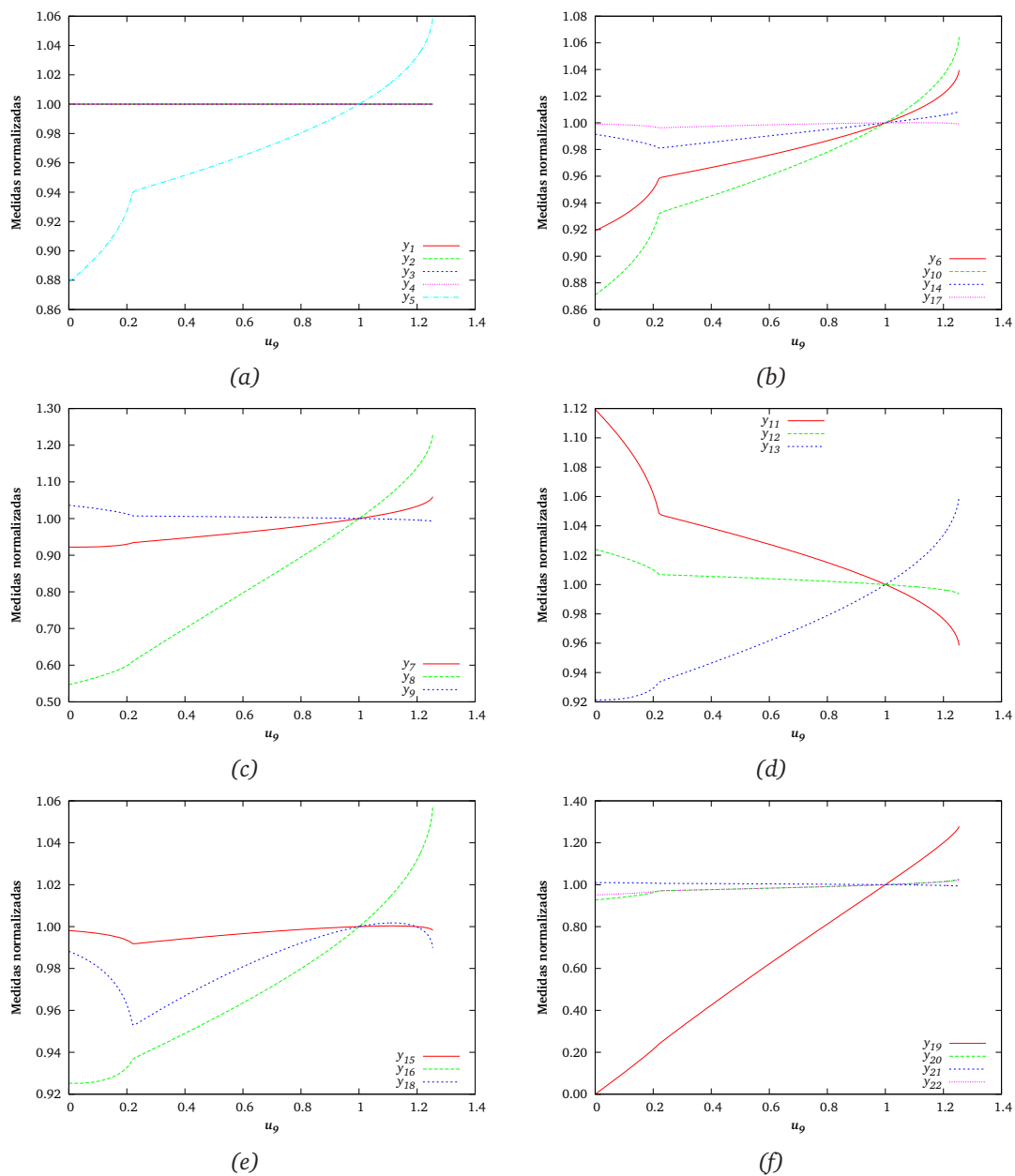


Figura B.7 Variáveis medidas em função do caudal de vapor no *stripper* ( $u_9$ ).

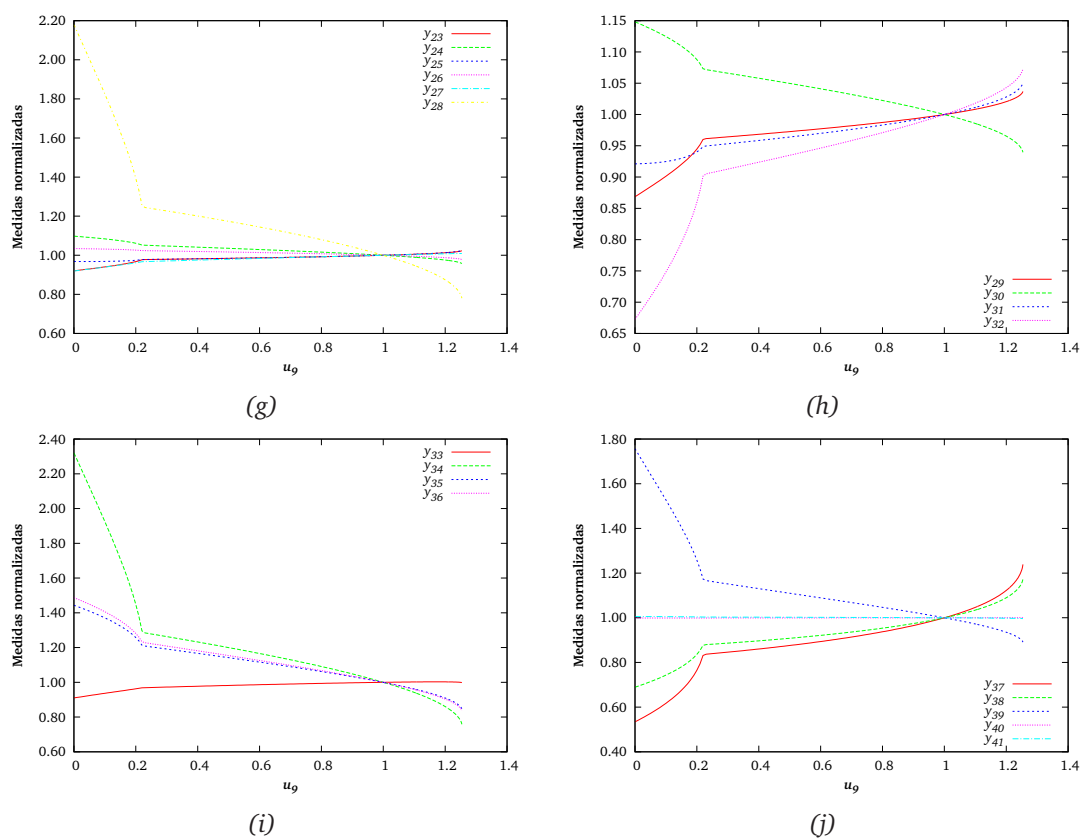
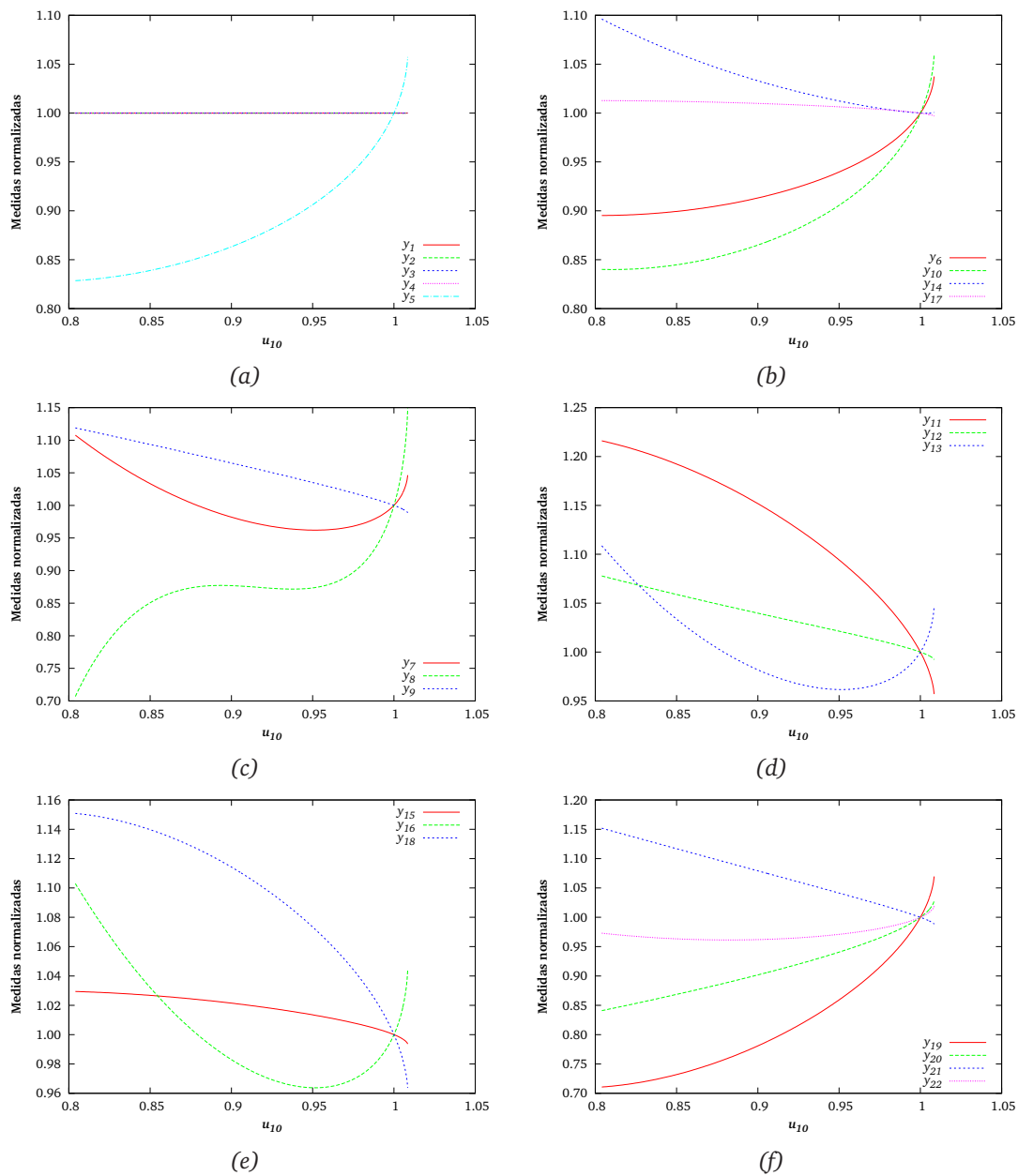
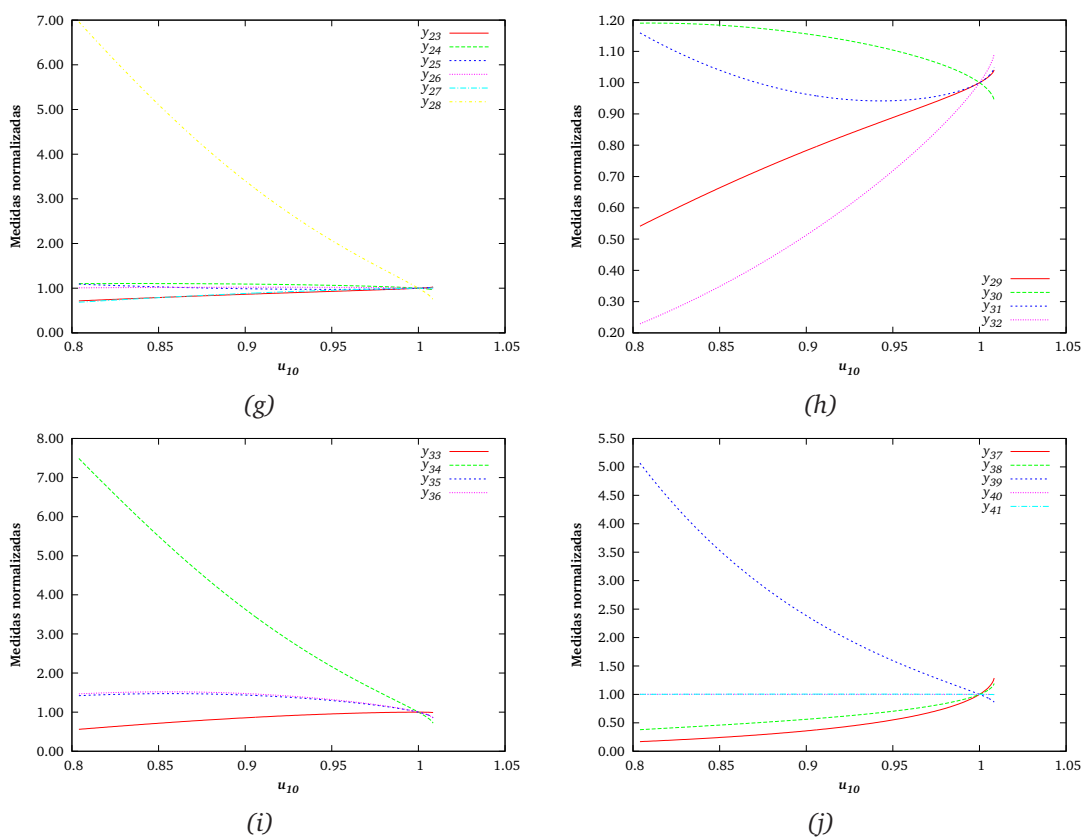


Figura B.7 (cont.) Variáveis medidas em função do caudal de vapor no *stripper* ( $u_0$ ).

Na Figura B.8 são apresentados os resultados relativos ao caudal da água de arrefecimento do reactor,  $u_{10}$ .

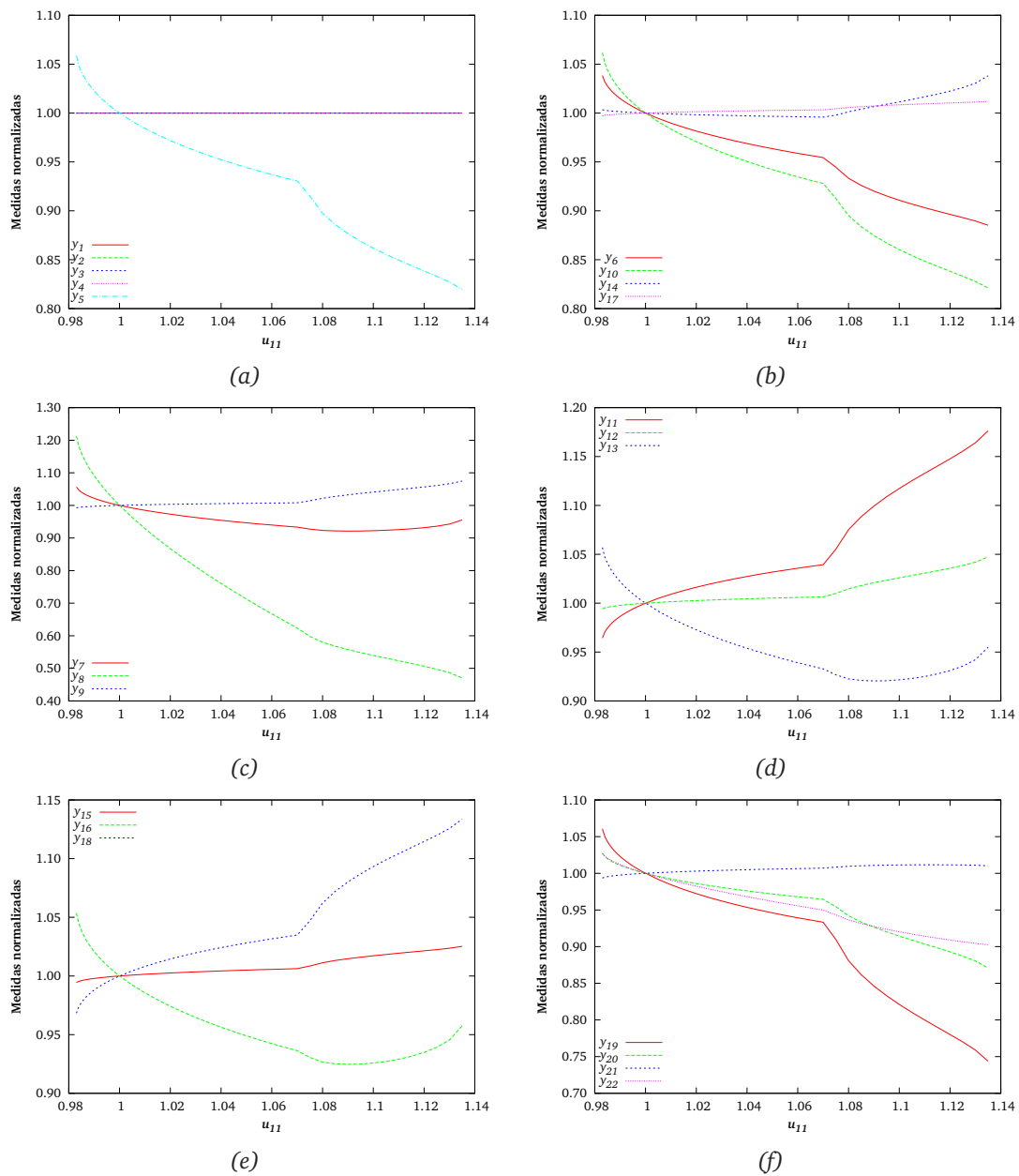


**Figura B.8** Variáveis medidas em função do caudal de água de arrefecimento do reactor ( $u_{10}$ ).

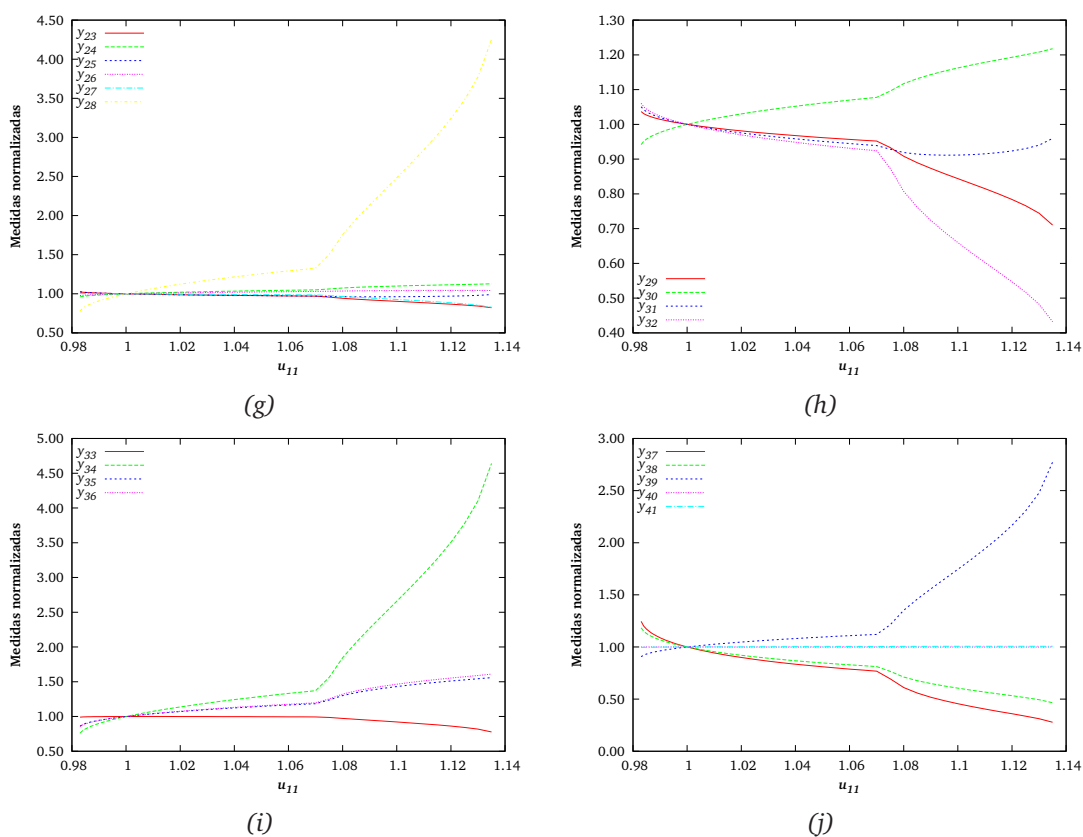


**Figura B.8 (cont.)** Variáveis medidas em função do caudal de água de arrefecimento do reactor ( $u_{10}$ ).

Na Figura B.9 são apresentados os resultados obtidos para o caudal da água de arrefecimento do separador,  $u_{11}$ .



**Figura B.9** Variáveis medidas em função do caudal de água de arrefecimento do condensador ( $u_{11}$ ).



**Figura B.9 (cont.)** Variáveis medidas em função do caudal de água de arrefecimento do condensador ( $u_{11}$ ).

Na Figura B.10 são apresentados os resultados para a velocidade de agitação do reator.

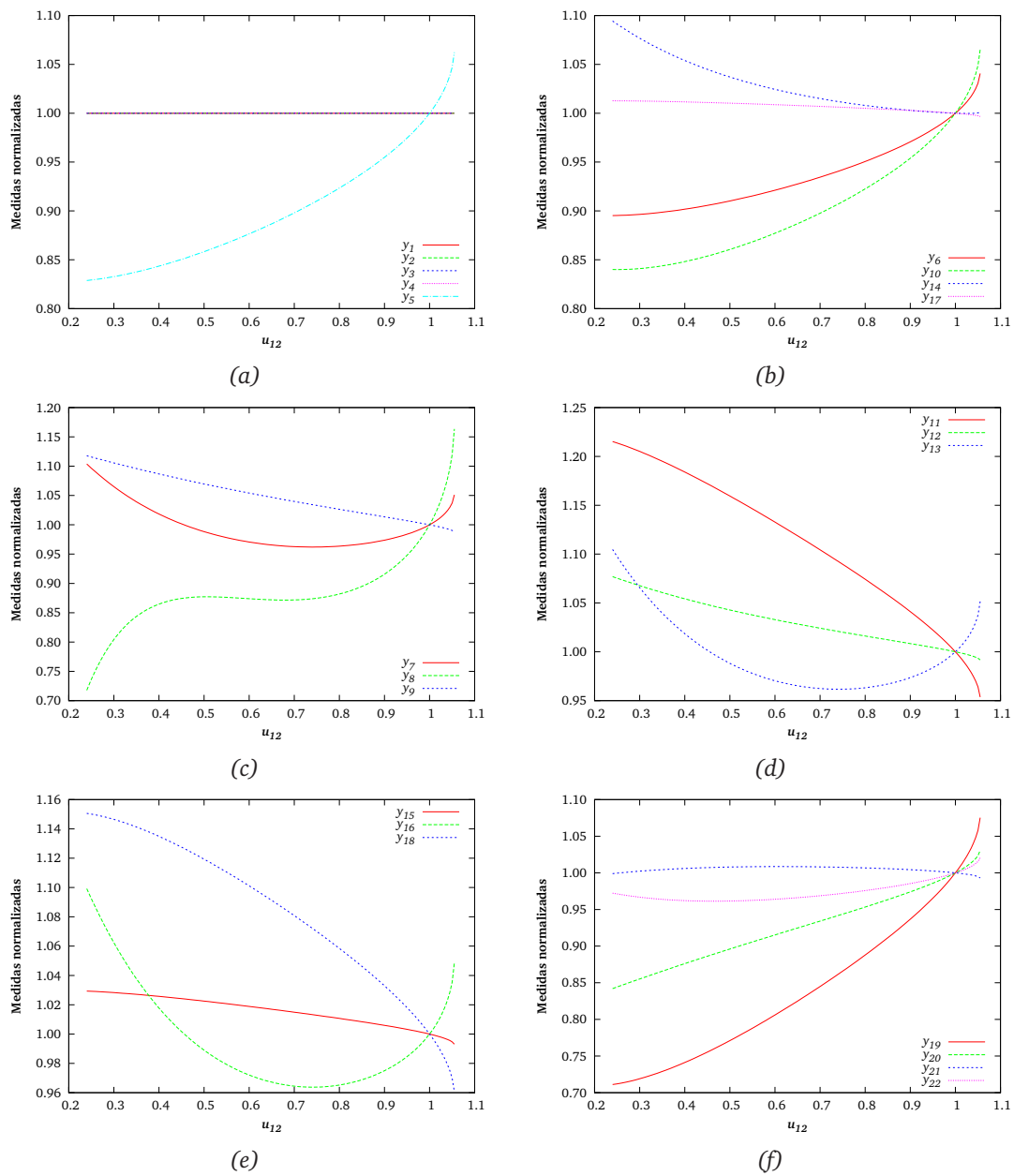


Figura B.10 Variáveis medidas em função da velocidade de agitação do reator ( $u_{12}$ ).



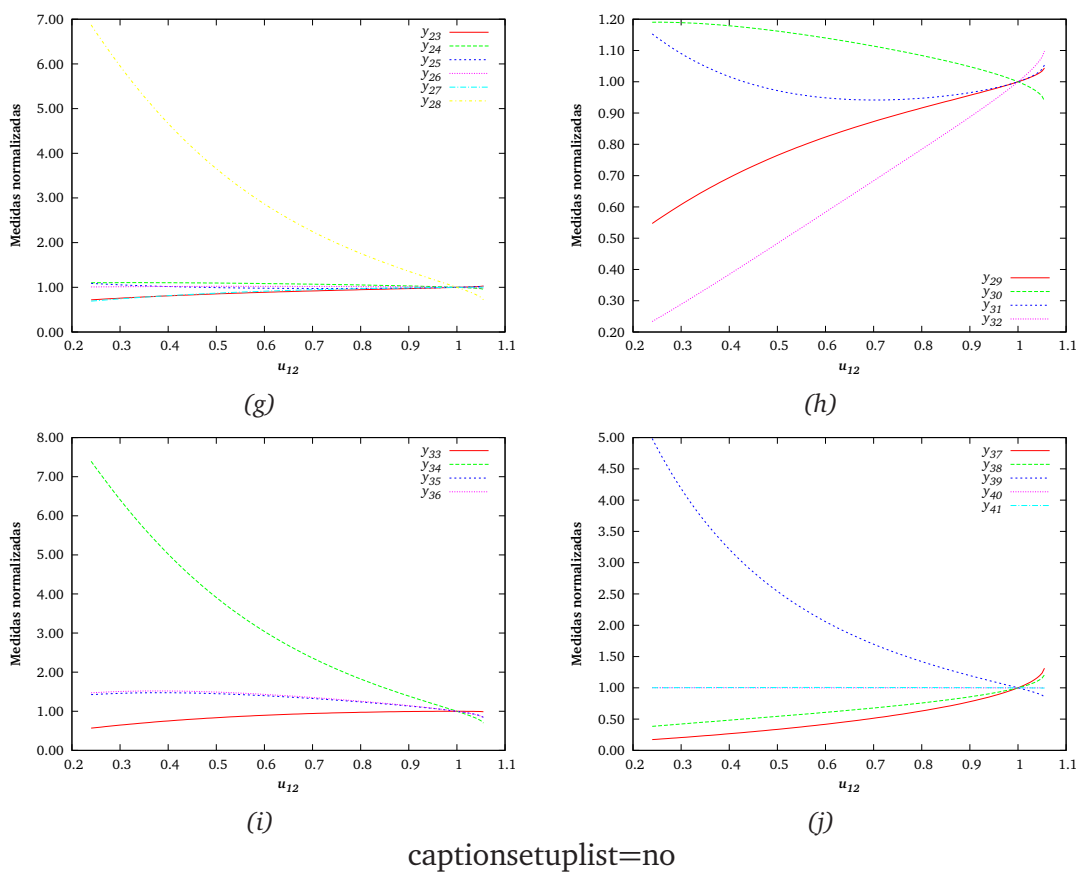


Figura B.10 (cont.) Variáveis medidas em função da velocidade de agitação do reactor ( $u_{12}$ ).



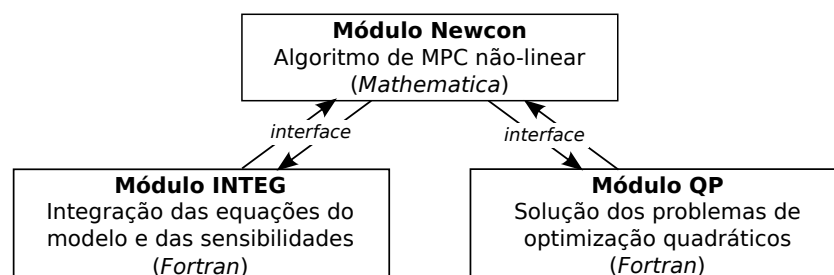
# Apêndice C

## Cálculo das sensibilidades

Neste Capítulo serão abordados diversos aspectos que influenciam a eficiência do cálculo das sensibilidades em modelos descritos por equações diferenciais, conforme descrito na Secção 2.4. São comparados os desempenhos de diferentes métodos de cálculo, de abordagens distintas à solução dos sistemas de equações lineares resolvidos e do efeito da tolerância especificada para o erro no cálculo destas sensibilidades. A necessidade desta análise é justificada pelos resultados obtidos durante o diagnóstico da eficiência computacional dos métodos utilizados durante a fase inicial deste trabalho.

### C.1 Implementação computacional

O algoritmo apresentado na Secção 4.1 foi implementado computacionalmente por Oliveira (1994), utilizando a estrutura modular representada na Figura C.1. Todos os resultados apresentados neste Capítulo são relativos a este algoritmo.

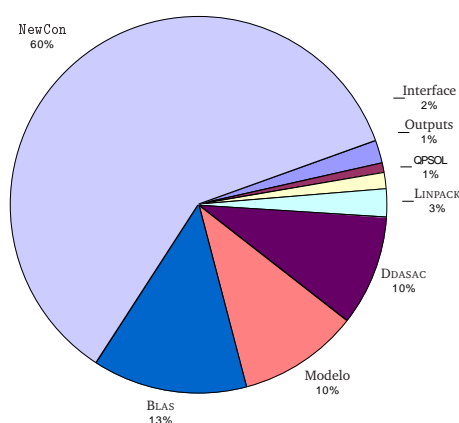


**Figura C.1** Representação esquemática da implementação do algoritmo de controle preditivo não-linear.

No módulo NewCon é implementado o algoritmo de controle, onde são resolvidos os problemas de otimização não-lineares, utilizando a linguagem de programação

incluída no programa MATHEMATICA (Wolfram, 1988). A decisão de implementar o algoritmo nesta linguagem deve-se à facilidade com que este *software* permite integrar informação numérica, simbólica e gráfica, e à sua natureza interactiva. No módulo *Integ* são integradas as equações do modelo e, se necessário, das respectivas sensibilidades. No módulo *QP* são resolvidos, em cada iteração do algoritmo, os problemas quadráticos. Ambos os módulos são escritos em FORTRAN utilizando, na implementação inicial, as rotinas DDASAC (Caracotsios e Stewart, 1985) e QPSOL (Gill et al., 1983). São também utilizadas as bibliotecas de álgebra linear BLAS (Lawson et al., 1979) e LINPACK (Anderson et al., 1999). A interface entre os diferentes módulos é feita através do protocolo MATHLINK (Wolfram, 1988). A estrutura modular tem a vantagem de ser relativamente simples a substituição das rotinas utilizadas em cada um dos módulos, tendo possibilitado a sua expansão de forma a tornar possível a sua aplicação a um mais alargado leque de problemas.

De modo a avaliar o desempenho computacional dos diferentes módulos, e das rotinas neles utilizadas, foi avaliada o tempo gasto em cada um deles na obtenção dos resultados de controlo apresentados na Secção 4.3.2. Na Figura C.2 é apresentada a distribuição do tempo de CPU para a obtenção dos resultados relativos ao controlo do processo na presença de variações aleatórias na cinéticas das reacções (apresentados nas Figuras 4.8 e 4.9).



**Figura C.2** Distribuição do tempo de CPU necessário à obtenção dos perfis apresentados nas Figuras 4.8 e 4.9, relativos ao controlo do processo TE com uma perturbação em degrau na temperatura de entrada da água de arrefecimento do reactor.

Na Figura C.2, as parcelas DDASAC, BLAS, LINPACK, Modelo e Interface pertencem ao módulo *Integ*. A parcela Modelo contabiliza o tempo gasto em todos os procedimentos necessários à avaliação das equações do modelo. Na parcela *Outputs* é contabilizado o tempo necessário para a obtenção das variáveis de saída a partir das variáveis de estado.

Como podemos verificar, cerca de 60% do tempo total é utilizado pelo módulo *NewCon*, o que se deve, em grande parte, à baixa eficiência numérica do MATHEMATICA, sobre-

tudo quando lida com problemas de grande dimensão, como neste caso. Também a forma como os diferentes procedimentos foram implementados não é particularmente adequada a este tipo de problemas. Um exemplo desta situação é a necessidade de formar explicitamente a totalidade da matriz dinâmica.

Relativamente à distribuição do tempo de CPU no módulo *Integ*, verifica-se que apenas 10% do tempo total é utilizado pela *DDASAC*, sendo igual percentagem utilizada para avaliação das equações do modelo e 16% utilizados na factorização e resolução de sistemas lineares. Tal indica um elevado número de avaliações das equações do modelo que, em grande parte, serão devidas ao facto de o Jacobiano ser obtido numericamente. Uma análise mais detalhada do tempo utilizado pelo módulo *Integ* revela que 47% do tempo utilizado por este módulo é devido ao cálculo das sensibilidades.

## C.2 Eficiência do cálculo das sensibilidades

A importância da eficiência do cálculo das sensibilidades no desempenho global do algoritmo de controlo é crucial. Analisando a formulação apresentada no Secção 4.1 podemos concluir que em cada iteração do algoritmo é necessário integrar  $p \cdot n_x \cdot (n_x + n_u)$  equações de sensibilidades e  $p \cdot n_x$  equações do modelo. No caso dos resultados apresentados na Secção 4.3.2, em que foi utilizando um horizonte de previsão  $p = 8$ , tal resulta em 24800 equações de sensibilidades e 400 equações do modelo, por iteração.

Os resultados atrás apresentados permitem concluir que o cálculo das sensibilidades não é feito de forma eficiente pela *DDASAC* em sistemas de grande dimensão. Esta rotina utiliza o método alternado directo para o cálculo das sensibilidade e diferenças finitas para obter os resíduos das equações das sensibilidades, métodos que têm algumas características indesejáveis quando aplicados a sistemas de grande dimensão. Nomeadamente a necessidade de calcular e factorizar a matriz  $A$  em cada passo do integrador pode representar um esforço de cálculo significativo, quer devido ao número de avaliações das equações do modelo, quando o Jacobiano é calculado numericamente, quer devido ao número de operações necessárias à factorização. A aplicação deste método ao controlo predictivo do processo TE confirma estas dificuldades uma vez que o perfil de tempo de CPU apresentado na Figura C.2 mostra que 47% do tempo utilizado no integrador é devido ao cálculo das sensibilidades.

Torna-se, por isso, importante considerar outros métodos de cálculo das sensibilidades, com características mais atractivas para aplicação a sistemas de grande dimensão, e testar a sua eficiência quando aplicados ao problema de controlo do processo TE.

### C.2.1 Métodos para o cálculo das sensibilidades

De forma a testar métodos alternativos à implementação do método alternado directo na DDASAC, esta foi substituída pela rotina DASPK. Esta rotina implementa os três métodos para o cálculo das sensibilidades apresentados na Secção 2.4 (incluindo o método alternado directo, com uma implementação ligeiramente diferente da utilizada na DDASAC), e permite a utilização de derivadas direccionais no cálculo dos resíduos das equações das sensibilidades. Uma característica importante desta rotina é o facto da matriz do Jacobiano (2.11) ser utilizada, depois de calculada e factorizada, o número máximo de vezes possível na solução das DAE's. Esta característica é particularmente vantajosa quando são utilizados os métodos alternado corrector ou simultâneo corrector no cálculo das sensibilidades, uma vez que estes reutilizam a matriz do Jacobiano utilizada para obter a solução das DAE's. Também importante é o facto da rotina ter sido concebida para sistemas de grande dimensão, sendo os algoritmos implementados de forma a minimizar os requisitos de memória e de armazenamento. Uma descrição mais detalhada da DASPK pode ser encontrada em [Li e Petzold \(1999\)](#).

Pelo atrás exposto, será de esperar uma melhoria da eficiência no cálculo das sensibilidades pela utilização da DASPK, pelo que se passou à sua utilização no módulo Integ. Na Tabela C.1 são apresentados os tempos de CPU utilizado pelo módulo Integ, o número de avaliações das equações do modelo e da matriz do Jacobiano necessárias para a simulação de 0.1 h (duas iterações do algoritmo de MPC) do processo TE, sem perturbações. É também apresentada a percentagem do tempo de CPU devido exclusivamente ao cálculo das sensibilidades. Os resultados apresentados foram obtidos utilizando tolerâncias absolutas e relativas de  $0.5 \times 10^{-7}$  para a convergência das variáveis de estado e de  $2.5 \times 10^{-7}$  para a convergência das sensibilidades.

**Tabela C.1** Tempo de CPU utilizado pelo módulo Integ, percentagem deste devido às sensibilidades, número de avaliações das equações do modelo,  $F$ , e da matriz do Jacobiano,  $J$ , necessários para a simulação de 0.1 h do processo TE, sem perturbações, utilizando como integrador as rotinas DDASAC e DASPK (método simultâneo corrector).

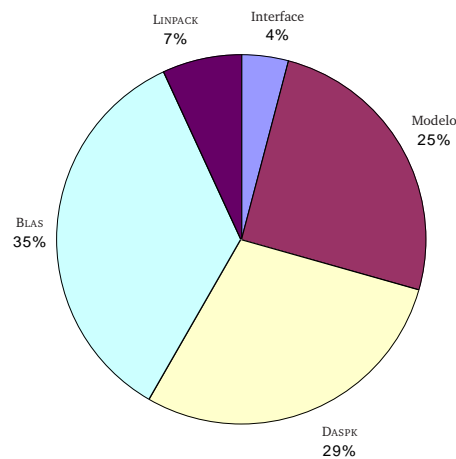
	DDASAC	DASPK
Tempo de CPU (s)	35.901	28.488
Sensibilidades (%)	50	24
Avaliações de $f$	292093	46143
Avaliações de $J$	5001	908

Como podemos verificar, é conseguida uma redução de 20% no tempo de CPU necessário para a integração e cálculo das sensibilidades através da utilização do método simultâneo corrector implementado pela DASPK. Esta diminuição é, em grande parte,

devida à diminuição de 84% do número de avaliações das equações do modelo necessárias, conseguida através do reaproveitamento da matriz do Jacobiano utilizada na solução das equações do modelo para o cálculo das sensibilidades.

### C.2.2 Solução de sistemas de equações lineares

Na Figura C.3 é apresentada a distribuição do tempo de CPU utilizado no módulo Integ pelas diferentes rotinas que o constituem. Os resultados apresentados foram obtidos utilizando o método simultâneo corrector, implementado pela DASPK, para o cálculo das sensibilidades. O tempo total de CPU é o apresentado na Tabela C.1.



**Figura C.3** Distribuição, no módulo Integ do tempo de CPU necessário para a simulação de 0.1 h do processo TE, sem perturbações, utilizando o método simultâneo corrector (DASPK).

Como podemos verificar a factorização e solução de sistemas de equações lineares, efectuadas pelas bibliotecas BLAS e LINPACK, são responsáveis por cerca de 42% do tempo total de CPU. Nenhuma destas bibliotecas implementa algoritmos otimizados para sistemas de média e grande dimensão ou com capacidade para aproveitar a estrutura dos problemas. De forma a conseguir possíveis melhorias na eficiência com que estas tarefas são executadas, os sistemas de equações lineares e a sua resolução foram estudados mais aprofundadamente.

Todos os sistemas de equações lineares resolvidos durante a integração do modelo e do cálculo das sensibilidades envolvem a matriz do Jacobiano das equações do modelo (2.11), pelo que o aproveitamento da sua estrutura pode introduzir melhorias na eficiência da resolução do problema. Na Figura C.4 é representada a estrutura de  $J$  para o processo TE nas condições correspondentes ao modo base de operação. Os elementos nulos representam cerca de 40% do número total de elementos.

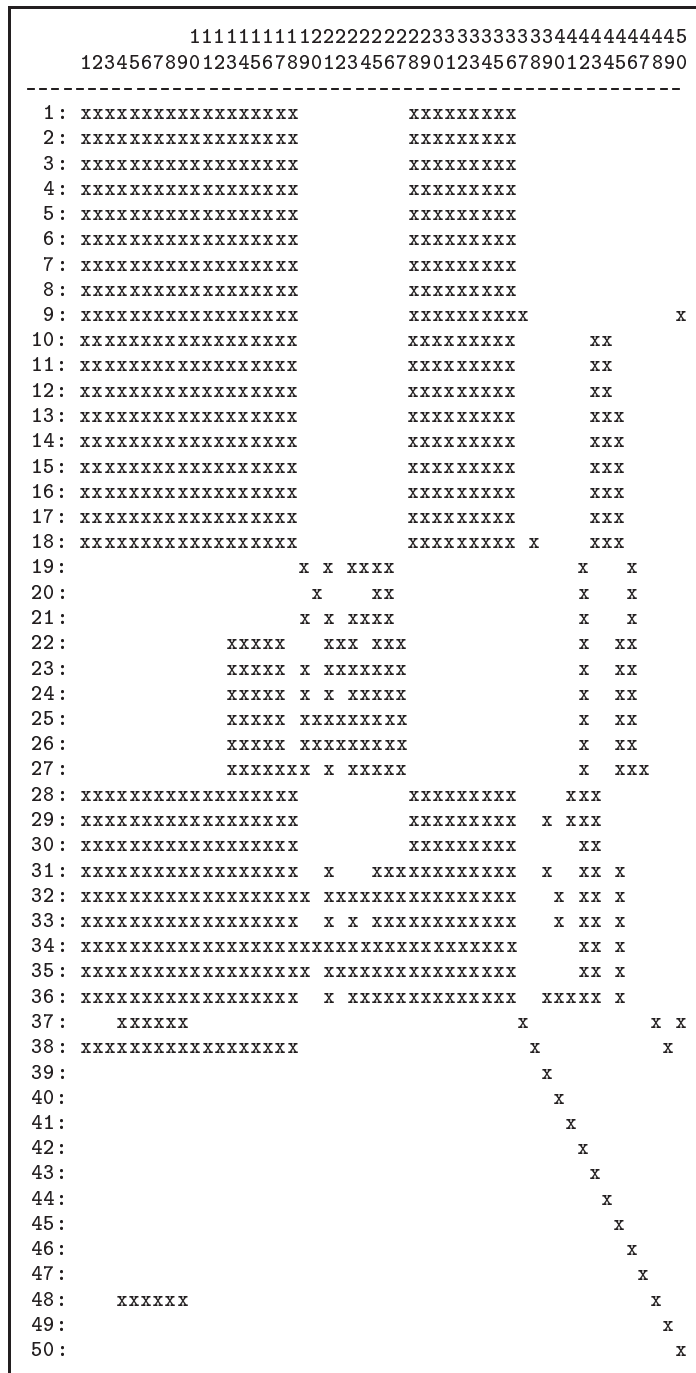


Figura C.4 Estrutura do Jacobiano do modelo do processo TE, incluindo o compensador, no estado estacionário correspondente ao modo base de operação.



Podemos observar a existência de alguma estrutura na matriz  $J$  do modelo em estado estacionário e, por análise do modelo do processo, podemos concluir que esta se mantém praticamente inalterada para diferentes condições de operação. Esta estrutura pode potencialmente ser aproveitada para, antes de qualquer operação com a matriz e apenas através de permutações das suas linhas e colunas, obter uma estrutura com características bem definidas, tal como diagonal por blocos ou *bordered block*. Estas estruturas são depois utilizadas por algoritmos de resolução de sistemas de equações lineares de forma a minimizar o número de operações necessárias para obter a solução. A generalidade dos métodos existentes para o rearranjo de matrizes é baseado na teoria dos grafos, nomeadamente na partição de grafos, uma vez que qualquer matriz pode ser descrita como um grafo orientado em que os nodos correspondem ao elementos não-nulos da matriz. Para rearranjar a matriz do Jacobiano do TE foram testados os algoritmos implementados nas bibliotecas SPARSKIT (Saad, 1994) e METIS (Karypis e Kumar, 1999). Nenhum dos algoritmos consegue melhorar a estrutura inicial, sendo esta devolvida como solução final.

Como alternativa ao reordenamento da matriz do Jacobiano *a priori* existem algoritmos que, durante a factorização, reordenam as linhas e colunas da matriz com o objectivo de minimizar o número de operações necessárias, quer durante a factorização quer durante a resolução do sistema de equações. Um destes algoritmos é implementado na rotina MA48 (Duff e Reid, 1993), que utiliza o método de Markowitz com algumas modificações. Na Tabela C.2 são apresentados os tempos de CPU utilizados no módulo `Integ` para a simulação de 0.1 h do processo TE, sem perturbações, utilizando para resolver os sistemas de equações lineares a biblioteca LINPACK e a rotina MA48, e para o cálculo das sensibilidades os métodos simultâneo corrector e alternado corrector.

**Tabela C.2** Tempo de CPU necessário para integração das equações do modelo e das sensibilidades na simulação de 0.1 h do processo TE utilizando diferentes métodos de solução de sistemas lineares, para diferentes métodos de cálculo das sensibilidades.

Sensibilidades	Tempo de CPU (s)	
	LINPACK	MA48
Simultâneo corrector	43.819	46.639
Alternado corrector	45.153	46.025

Como podemos verificar os tempos de CPU necessários para simulação utilizando a MA48 são superiores aos necessários quando é utilizada a LINPACK. Uma vez que todas as quantidades que influenciam o tempo de CPU (número de passos do integrador, avaliações do modelo e do Jacobiano) se mantêm inalterados, podemos concluir que o aumento do tempo de CPU apenas é devido ao maior processamento da matriz feito pela MA48.

Os resultados atrás apresentados mostram que a estrutura e a esparsidade da matriz do Jacobiano do modelo do processo TE não são suficientes para serem explorado eficientemente pelos algoritmos testados. Apesar de estes métodos terem sido desenvolvidos para lidar com sistemas de equações lineares de dimensões elevadas, são sobretudo eficientes quando a percentagem de elementos não-nulos é bastante reduzida e significativamente inferior à do problema em estudo.

### C.2.3 Eficiência das rotinas de álgebra linear

Os resultados apresentados permitem concluir que a forma mais eficiente de resolver os sistemas de equações lineares necessários à integração das equações do modelo e das sensibilidades do modelo do processo TE é tratando as matrizes dos sistemas como densas. Tal é feito utilizando a biblioteca LINPACK que utiliza extensamente as rotinas da biblioteca BLAS. O elevado número de sistemas de equações lineares que é resolvido em cada iteração do algoritmo de controlo tem como consequência um elevado número de chamadas das rotinas da BLAS. Por exemplo, para obter os resultados apresentados na Tabela C.2 (duas iterações do algoritmo de controlo) as rotinas da BLAS são chamadas 44,166,749 vezes quando é utilizado o método simultâneo corrector e 45,057,816 vezes quando é utilizado o método alternado corrector. Daqui, e da distribuição do tempo de CPU apresentada na Figura C.3, facilmente se conclui que a utilização de subrotinas de álgebra linear mais eficazes pode permitir ganhos significativos na eficiência da resolução do problema.

Uma forma de obter rotinas de álgebra linear mais eficazes é otimizar as rotinas da BLAS, nomeadamente a forma como as suas operações são efectuadas e como é feito o acesso à memória do computador. Esta optimização é um processo não-trivial e altamente dependente das características do computador utilizado, tais como a arquitectura do processador e o tipo e quantidade de memória. Whaley et al. (2000) propõem um processo, empírico, de optimização automática do desempenho das rotinas da BLAS para diferentes configurações de *hardware*. O procedimento proposto permite a construção de uma biblioteca, a ATLAS, que contém as mesmas rotinas da BLAS, mas optimizadas para a configuração específica de *hardware* em que são criadas. Na Tabela C.3 são apresentados os tempos de CPU necessários para a integração das equações do modelo e das sensibilidades na simulação de 0.1 h do processo TE utilizando a biblioteca BLAS, sem optimização, e a biblioteca ATLAS, optimizada. Os resultados são apresentados para os métodos simultâneo corrector e alternado corrector, implementados na DASPK.

Como podemos verificar a utilização de rotinas de álgebra linear optimizadas, ATLAS, permite uma redução de aproximadamente 10% no tempo de CPU total necessário para a integração das equações, para ambos os métodos de cálculo das sensibilidades. Esta diminuição é devida a uma redução de cerca de 30% no tempo utilizado nas

**Tabela C.3** Tempo de CPU, em segundos, necessário para a integração das equações do modelo e das sensibilidades na simulação de 0.1 h do processo TE, sem perturbações, utilizando as rotinas de álgebra linear com e sem otimização, para diferentes métodos de cálculo das sensibilidades.

	Sem otimização (s)		Com otimização (s)	
	Total	BLAS	Total	ATLAS
simultâneo corrector	43.819	15.36	39.556	10.50
alternado corrector	45.153	15.81	40.100	10.77

rotinas da BLAS, uma vez que o tempo gasto nas restantes rotinas necessárias à integração se mantém inalterado nas duas situações. A percentagem de tempo de CPU utilizado pelas rotinas de álgebra linear é reduzido de 35% para 26%. Os resultados apresentados foram obtidos numa Digital Personal Workstation 433au mas, dos resultados apresentados por [Whaley et al. \(2000\)](#), são de esperar reduções semelhantes noutras arquitecturas.

#### C.2.4 Tolerâncias no cálculo das sensibilidades

Como já foi referido, nas situações em que não é possível o cálculo analítico das derivadas  $\partial F / \partial \theta_i$  é necessário recorrer a aproximações numéricas destas. A utilização destas aproximações implica a introdução de erros, quer devido à aproximação em si, quer devido à precisão finita com que os cálculos são efectuados. Quando o método utilizado são as diferenças finitas *forward* com um intervalo  $\delta$ ,

$$\frac{\partial F}{\partial \theta_i} = \frac{F(\dot{v}, v, \theta + \delta_i e_i, t) - F(\dot{v}, v, \theta, t)}{\delta_i},$$

[Gill et al. \(1981\)](#) mostram que o erro cometido na aproximação é limitado por

$$\frac{\delta_i}{2} \left| \frac{\partial^2 F(\dot{v}, v, \xi, t)}{\partial \theta_i^2} \right| + \frac{2}{\delta_i} \epsilon_A, \quad (\text{C.1})$$

onde  $\xi$  é um vector igual a  $\theta$  à excepção do elemento  $i$ ,  $\xi_i$ , que pertence ao intervalo  $[\theta_i, \theta_i + \delta_i]$ , e  $\epsilon_A$  é o limite do erro no cálculo de  $F$  em  $\theta$  e  $\theta + \delta_i e_i$ . O intervalo  $\delta_{i,F}$  que minimiza esta soma é

$$\delta_{i,F} = 2 \sqrt{\frac{\epsilon_A}{\left| \frac{\partial^2 \theta_i F(\dot{v}, v, \xi, t)}{\partial \theta_i^2} \right|}}.$$

Substituindo em (C.1), verificamos que o limite mínimo do erro é

$$2\sqrt{\epsilon_A \left| \frac{\partial^2 F(\dot{v}, v, \xi, t)}{\partial \theta_i^2} \right|}.$$

Uma vez que  $\xi_i$  não é conhecido, são necessárias algumas suposições de forma a obter uma estimativa do erro. Considerando

$$\epsilon_A = |F| \epsilon_R,$$

onde  $\epsilon_R$  é o limite do erro relativo no cálculo de  $F$  em  $\theta$  e  $\theta + \delta_i e_i$ , e se

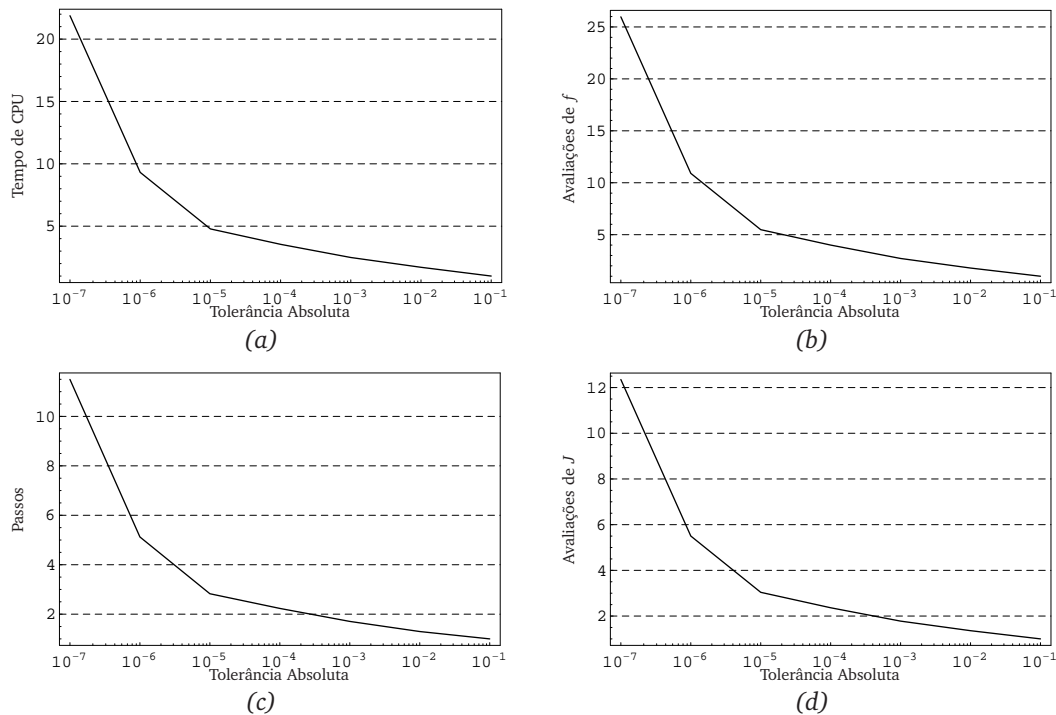
$$O(|F|) = O\left(\left|\frac{\partial F}{\partial \theta_i}\right|\right) = O\left(\left|\frac{\partial^2 F}{\partial \theta_i^2}\right|\right)$$

em todos os pontos entre  $[\theta_i, \theta_i + \delta_i]$ , conclui-se que o erro relativo na aproximação a  $\partial F / \partial \theta_i$  por diferenças finitas *forward* é de ordem  $\sqrt{\epsilon_R}$ . Tal apenas é válido se forem verificadas todas as suposições feitas.

Da análise apresentada, ainda que apenas válida para diferenças finitas *forward*, facilmente podemos concluir que o erro cometido na aproximação a  $\partial F / \partial \theta_i$  por qualquer outro método depende do erro cometido no cálculo das variáveis de estado,  $v$ . No entanto a relação entre ambos os erros é difícil de determinar, uma vez que para tal seria necessário conhecer qual a influência do erros em  $v$  no valor de  $F$ , ou seja, conhecer  $\epsilon_R$ , no caso da aproximação por diferenças finitas. No caso da aproximação por derivadas direccionais acresce ainda a dificuldade de não serem conhecidos resultados concretos quanto à sua precisão, sendo o intervalo utilizado na DASPCK calculado através de fórmulas empíricas.

As relações atrás apresentadas mostram claramente a existência de uma relação entre o erro cometido na integração das equações do modelo, para calcular as variáveis de estado, e o erro cometido no cálculo das sensibilidades. Os perfis de tempo de CPU apresentados nas Figura C.2 e C.3 foram obtidos utilizando tolerâncias absolutas e relativas de  $0.5 \times 10^{-7}$  para a convergência das variáveis de estado e de  $2.5 \times 10^{-7}$  para a convergência das sensibilidades. Nas Figuras C.5, C.6 e C.7 são apresentadas as variações, com a tolerância absoluta para a convergência das sensibilidades, do esforço computacional para a simulação de 0.1 h do processo TE, sem perturbações, utilizando como integrador as rotinas DDASAC e DASPCK (métodos alternado corrector e simultâneo corrector). Em todas as simulações as tolerâncias absoluta e relativa para a convergência das variáveis de estado são de  $0.5 \times 10^{-7}$ .

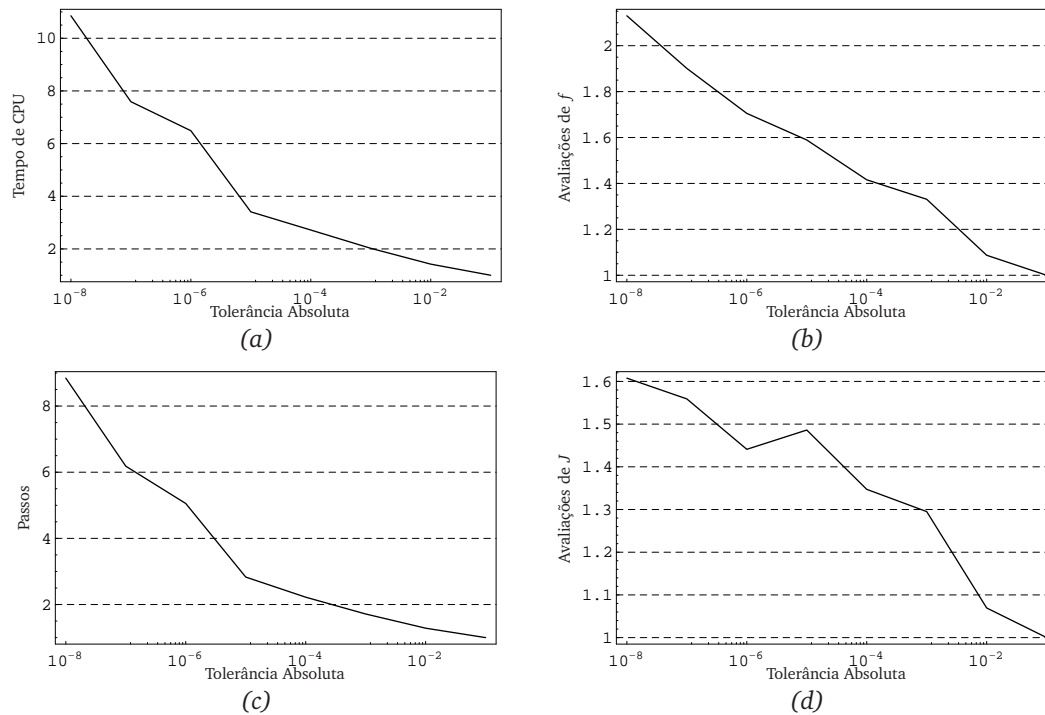
Como podemos verificar, para as três situações apresentadas, o esforço computacional necessário para a simulação, medido pelo tempo de CPU, cresce muito rapidamente para tolerâncias inferiores a  $10^{-5}$ . Tal permite concluir que não é possível calcular as



**Figura C.5** Variações normalizadas, com o logaritmo da tolerância absoluta imposta para a convergência das sensibilidades, do tempo de CPU utilizado na integração do modelo e das sensibilidades (a), do número de avaliações das equações do modelo (b), do número de passos do integrador (c) e do número de avaliações do Jacobiano (d) necessários para a simulação de 0.1h do processo TE utilizando a DDASAC como integrador.

sensibilidades com um erro inferior a  $10^{-5}$  sem que seja necessário calcular as variáveis de estado com um erro muito inferior à tolerância imposta (aproximadamente  $10^{-7}$ , uma vez que todas as variáveis de estado são normalizadas), confirmando os resultados da análise atrás apresentada. Para que tal seja possível é necessário um aumento significativo do número de passos dados pelo integrador, o que implica directamente um igualmente significativo aumento do número de sistemas de equações lineares que é necessário resolver. No caso da DDASAC acresce ainda o facto de, para cada passo do integrador, ser necessário o cálculo e a factorização da matriz do Jacobiano, pelo que o aumento do tempo de CPU é muito superior ao observado com a DASPK. Nas Figuras C.6 e C.7 podemos confirmar que o número de avaliações do Jacobiano, quando são utilizados os métodos simultâneo corrector e alternado corrector, não cresce de forma tão acentuada como o número de passos, com a diminuição da tolerância. Tal deve-se ao facto de nestes métodos ser possível a reutilização da matriz do Jacobiano utilizada no cálculo das variáveis de estado, resultando deste facto um menor aumento do tempo de CPU.

A partir dos resultados apresentados podemos concluir que para tolerâncias inferiores a  $10^{-5}$  o esforço computacional necessário é sobretudo devido ao cálculo das sensi-

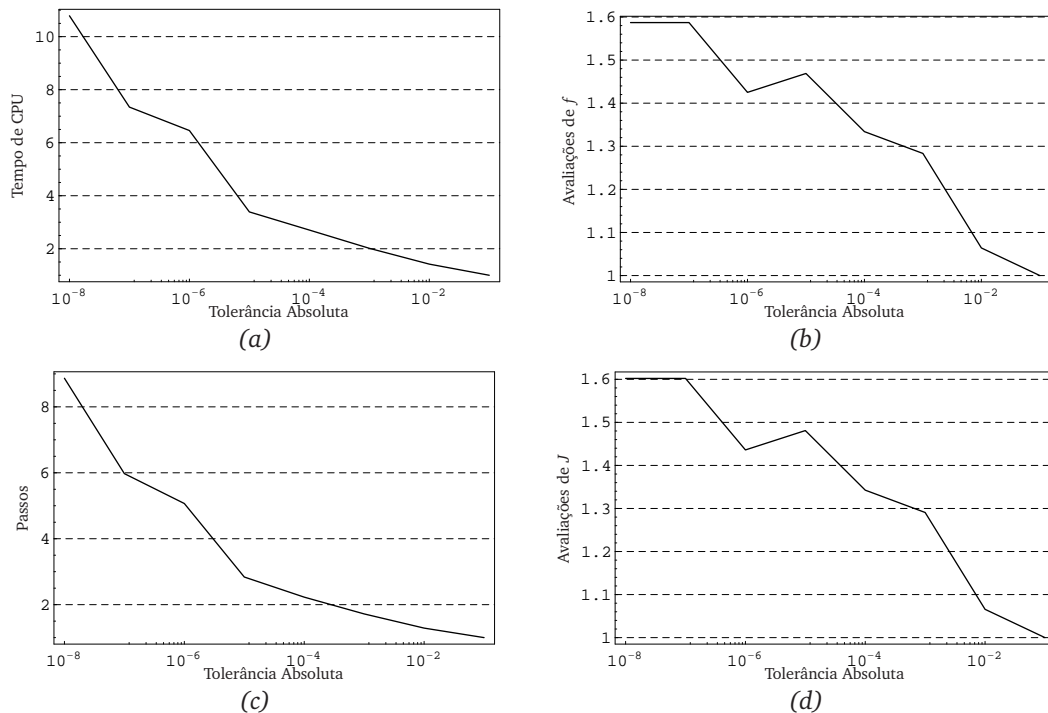


**Figura C.6** Variações normalizadas, com o logaritmo da tolerância absoluta imposta para a convergência das sensibilidades, do tempo de CPU utilizado na integração do modelo e das sensibilidades (a), do número de avaliações das equações do modelo (b), do número de passos do integrador (c) e do número de avaliações do Jacobiano (d) necessários para a simulação de 0.1h do processo TE utilizando a DASPK (método alternado corrector) como integrador.

bilidades, quando os estados são calculados com uma tolerância igual ou inferior, no caso do processo TE. Para outros modelos facilmente se conclui a existência de relações semelhantes entre os erros no cálculo dos estados e das sensibilidades. Logo é necessária alguma precaução na escolha das tolerâncias utilizadas uma vez estas podem ser determinantes no esforço computacional necessário.

### C.3 Comparação dos métodos de cálculo da sensibilidades

Nesta Secção são apresentadas comparações entre a *performance* dos diferentes métodos apresentados para o cálculo das sensibilidades e as diferentes estratégias de resolução de sistemas de equações lineares. Todas as comparações são feitas em função da tolerância imposta no cálculo das sensibilidades uma vez que, como foi mostrado na Secção anterior, esta tem uma influência determinante no esforço computacional necessário para a simulação. Todos os resultados apresentados são relativos a simulações de 0.1 h de controlo do processo TE, sem perturbações, utilizando a estrutura de



**Figura C.7** Variações normalizadas, com o logaritmo da tolerância absoluta imposta para a convergência das sensibilidades, do tempo de CPU utilizado na integração do modelo e das sensibilidades (a), do número de avaliações das equações do modelo (b), do número de passos do integrador (c) e do número de avaliações do Jacobiano (d) necessários para a simulação de 0.1h do processo TE utilizando a DASPK (método simultâneo corrector) como integrador.

controlo proposta na Capítulo 3. As tolerâncias absoluta e relativa para o cálculo das variáveis de estados são de  $0.5 \times 10^{-7}$  em todas as situações. A matriz do Jacobiano é obtida numericamente por diferenças finitas *forward*, uma vez que o seu cálculo analítico não é possível.

Na Tabela C.4 são apresentados os resultados relativos à eficiência do método alternado corrector, implementado pela DASPK. Os diferentes tempos de CPU referem-se a três formas distintas de resolver sistemas de equações lineares, nomeadamente a utilização das bibliotecas LINPACK e BLAS, das rotinas MA48 e das bibliotecas LINPACK e ATLAS. O número de passos do integrador e de avaliações da matriz do Jacobiano são os mesmos nas três situações.

Como podemos observar, o número de passos do integrador cresce muito mais rapidamente para tolerâncias inferiores a  $10^{-5}$  do que para tolerâncias superiores a este valor. Este aumento não se reflecte no número de avaliações da matriz do Jacobiano uma vez que no método alternado corrector, sendo reaproveitada, não é necessário o seu cálculo em cada passo. Na Figura C.8 são representados graficamente os valores da Tabela C.4, utilizando as razões entre os diferentes tempos de CPU. Nomeadamente são apresentadas as razões entre os tempos de CPU utilizando a biblioteca



**Tabela C.4** Número de passos do integrador, de avaliações da matriz do Jacobiano e tempos de CPU utilizando diferentes estratégias de solução de sistemas lineares, necessários para a simulação do processo TE, utilizando o método alternado corrector no cálculo das sensibilidades (DASPK).

Tolerância	Passos	Avaliações $J$	Tempo CPU (s)		
			LINPACK <sup>a</sup>	MA48	ATLAS <sup>b</sup>
$10^{-1}$	1091	575	5.946	6.860	5.201
$10^{-2}$	1402	616	8.460	9.547	7.456
$10^{-3}$	1870	746	12.032	13.393	10.675
$10^{-4}$	2424	776	16.161	17.705	14.333
$10^{-5}$	3088	856	20.266	22.076	17.953
$10^{-6}$	5515	830	38.583	41.070	34.260
$10^{-7}$	6532	902	45.153	46.025	40.100
$10^{-8}$	9648	926	64.530	68.094	57.468

<sup>a</sup> utilizando a biblioteca BLAS

<sup>b</sup> utilizando a biblioteca LINPACK

LINPACK e as rotinas MA48, e utilizando a bibliotecas BLAS e ATLAS.

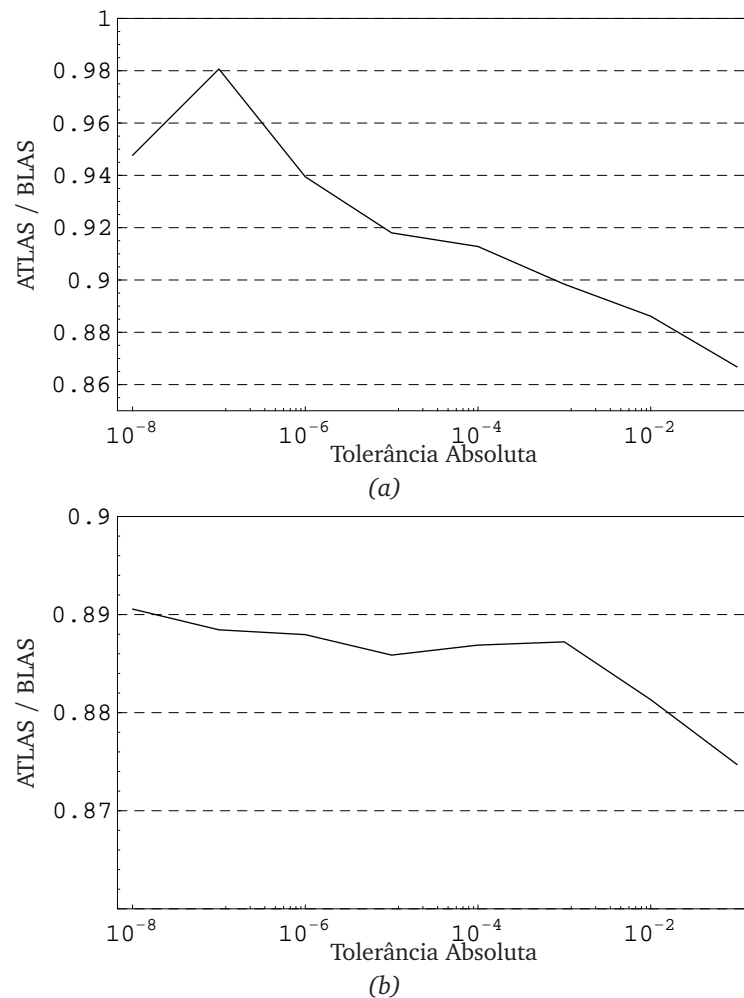
Como podemos verificar a utilização da LINPACK resulta sempre numa maior eficiência na factorização e resolução de sistemas de equações lineares, pelas razões já apresentadas na Secção C.2.2. Também podemos verificar que a utilização das rotinas de álgebra linear optimizadas, fornecidas pela ATLAS, é sempre significativamente mais eficiente que a utilização das mesmas rotinas não optimizadas, permitindo reduções do tempo de CPU de 10 a 12.5%.

Na Tabela C.5 são apresentados os resultados relativos à eficiência do método simultâneo corrector, implementado pela DASPK. São apresentados os tempos de CPU obtidos com a utilização da biblioteca LINPACK e das rotinas MA48 para a resolução de sistemas de equações lineares.

Mais uma vez observamos um aumento da taxa de crescimento do número de passos do integrador para tolerâncias inferiores a  $10^{-5}$  sem que se observe um crescimento semelhante no número de avaliações da matriz do Jacobiano. Relativamente ao tempo de CPU repete-se a situação observada para o método alternado corrector, ou seja, a utilização da LINPACK resulta numa eficiência superior à obtida com a utilização da MA48. Tal seria expectável, uma vez que a matriz do Jacobiano é a mesma nos dois métodos.

Na Figura C.9 é apresentada a evolução com a tolerância no cálculo das sensibilidades da razão entre os tempos de CPU utilizando os métodos alternado corrector e simultâneo corrector, ambos utilizando a LINPACK. Podemos verificar que não existem diferenças significativas entre os tempos de CPU necessários aos dois métodos para a





**Figura C.8** Razões entre os tempos de CPU necessários para a simulação utilizando o método alternado corrector e a biblioteca LINPACK e as rotinas MA48 (a), e as bibliotecas ATLAS e BLAS (b) calculadas a partir dos valores apresentados na Tabela C.4.

simulação.

Na Tabela C.6 são apresentados os resultados relativos à eficiência do método alternado directo, implementado pela DDASAC. São apresentados os tempos de CPU obtidos utilizando a biblioteca LINPACK e as bibliotecas BLAS e ATLAS.

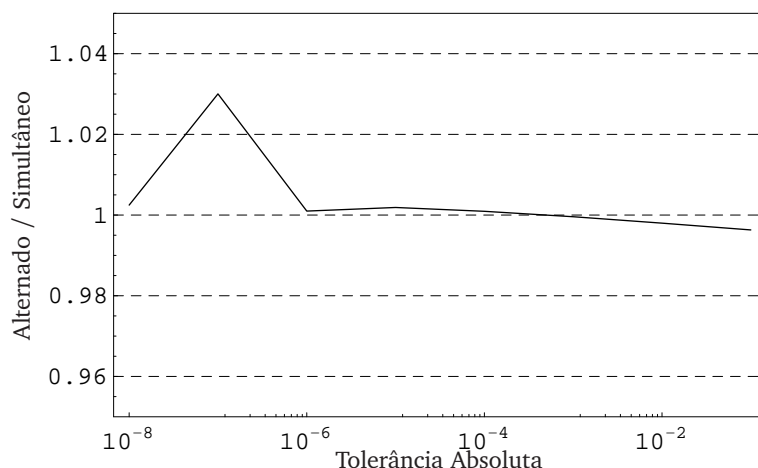
À semelhança dos outros métodos, também com o método alternado directo se observa um crescimento mais rápido do número de passos do integrador para tolerâncias do cálculo das sensibilidades inferiores a  $10^{-5}$ . No entanto, ao contrário dos outros métodos, este crescimento implica um igualmente rápido aumento do número de avaliações da matriz do Jacobiano, uma vez que é necessária a avaliação desta em cada passo do integrador. Como consequência verifica-se um maior aumento do tempo de CPU.

Para facilitar a comparação entre o método alternado corrector, implementado pela

**Tabela C.5** Número de passos do integrador, de avaliações da matriz do Jacobiano e tempos de CPU utilizando diferentes estratégias de solução de sistemas lineares, necessários para a simulação do processo TE, utilizando o método simultâneo corrector no cálculo das sensibilidades (DASPK).

Tolerância	Passos	Avaliações $J$	Tempo CPU (s)	
			LINPACK <sup>a</sup>	MA48 <sup>a</sup>
$10^{-1}$	1088	578	5.968	6.904
$10^{-2}$	1402	616	8.477	9.543
$10^{-3}$	1870	746	12.037	13.447
$10^{-4}$	2424	776	16.146	17.663
$10^{-5}$	3088	856	20.228	22.062
$10^{-6}$	5514	830	38.545	40.842
$10^{-7}$	6500	926	43.819	46.639
$10^{-8}$	9648	926	64.371	67.894

<sup>a</sup> utilizando a biblioteca BLAS



**Figura C.9** Razões entre os tempos de CPU necessários para a simulação utilizando os métodos alternado corrector e simultâneo corrector calculadas a partir dos valores apresentados nas Tabelas C.4 e C.5.

DASPK, e o método alternado directo, implementado pela DDASAC, alguns dos respectivos resultados relativos à eficiência são apresentados na Tabela C.7. Os resultados do método alternado corrector foram obtidos utilizando as bibliotecas LINPACK e ATLAS, e os do método alternado directo foram obtidos utilizando a LINPACK e a BLAS. Foram estes os resultados escolhidos por serem os melhores para cada um dos métodos. Na Figura C.10 são representadas graficamente as razões entre os valores obtidos para os dois métodos para as diferentes tolerâncias.

Os resultados apresentados permitem concluir que o esforço computacional necessário para a simulação é dominado pela resolução de sistemas de equações lineares, e

**Tabela C.6** Número de passos do integrador, de avaliações da matriz do Jacobiano e tempos de CPU utilizando diferentes estratégias de solução de sistemas lineares, necessários para a simulação do processo TE, utilizando o método alternado directo no cálculo das sensibilidades (DDASAC).

Tolerância	Passos	Avaliações $J$	Tempo CPU (s)	
			BLAS <sup>a</sup>	ATLAS
$10^{-1}$	860	802	3.326	3.844
$10^{-2}$	1116	1090	5.665	6.479
$10^{-3}$	1468	1426	8.276	9.699
$10^{-4}$	1920	1895	11.787	13.458
$10^{-5}$	2431	2437	15.899	18.162
$10^{-6}$	4406	4413	30.978	35.384
$10^{-7}$	9886	9906	72.759	83.047

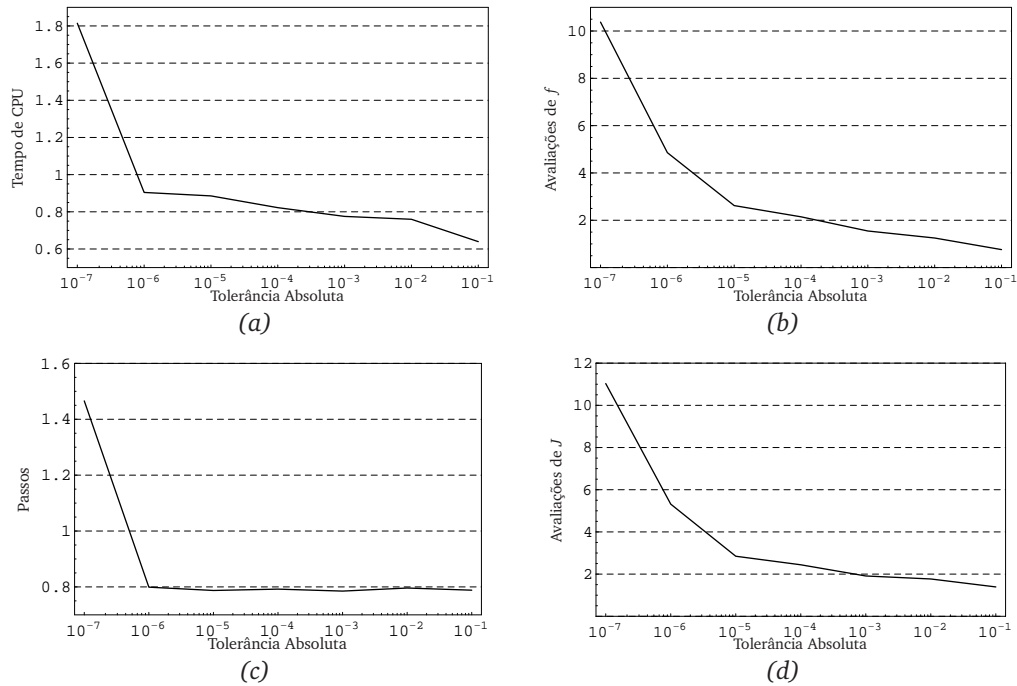
<sup>a</sup> utilizando a biblioteca LINPACK

**Tabela C.7** Comparação entre o método alternado corrector, DASPCK, e o método alternado directo, DDASAC.

Tolerância	Passos		Avaliações $J$		Tempo CPU (s)	
	DDASAC	DASPCK	DDASAC	DASPCK	DDASAC	DASPCK
$10^{-5}$	2431	3088	2437	856	15.899	17.953
$10^{-6}$	4406	5515	4413	830	30.978	34.260
$10^{-7}$	9886	6532	9906	902	72.759	40.100

não pela avaliação das equações do modelo necessárias ao cálculo da matriz do Jacobiano. Observando os resultados apresentados para as tolerâncias de  $10^{-5}$  e  $10^{-6}$  verificamos que o número de avaliações da matriz do Jacobiano necessário pela DDASAC é 3 e 5 vezes superior ao da DASPCK, respectivamente. No entanto, o número de passos da DDASAC é cerca de 20% inferior ao da DASPCK. Cada passo do integrador, quando são calculadas as sensibilidades, implica a resolução dos sistemas de equações lineares (2.10), até a convergência, e (2.12) quando é utilizada a DDASAC. Quando é utilizada a DASPCK é necessário resolver os sistemas (2.10) e (2.13), ambos até a convergência. Daqui resulta que o tempo de CPU utilizado pela DASPCK é cerca de 10% superior ao utilizado pela DDASAC, para ambas as tolerâncias. Esta conclusão é reforçada pelos resultados obtidos para o método alternado corrector para as tolerâncias de  $10^{-5}$  e  $10^{-6}$ . Entre estas tolerâncias verifica-se um aumento do número de passos do integrador, mas uma diminuição do número de avaliações da matriz do Jacobiano, resultando num aumento do tempo de CPU.

Para a tolerância de  $10^{-7}$ , quer o número de passos quer o número de avaliações da matriz do Jacobiano necessários à DDASAC são muito superiores ao necessários à DASPCK, o que inevitavelmente resulta num tempo de CPU muito superior. Este



**Figura C.10** Razões entre a DASPK (método alternado corrector, LINPACK e ATLAS) e a DDASAC (método alternado directo), calculadas a partir dos valores apresentados nas Tabelas C.4 e C.6

comportamento deve-se à dificuldade da DDASAC em calcular as sensibilidades com tolerâncias apertadas, não sendo possível o seu cálculo com tolerâncias inferiores a  $10^{-7}$ . Utilizando a DASPK é possível o cálculo das sensibilidades com tolerâncias até  $10^{-9}$ , para os métodos alterando corrector e simultâneo corrector, apesar do esforço computacional necessário ser extremamente elevado.

Como podemos verificar, pelos valores das Tabelas C.4, C.5 e C.6, para tolerâncias superiores a  $10^{-6}$  o método mais eficiente para calcular as sensibilidades do processo TE necessárias ao algoritmo de controlo é o método alternado directo, implementado pela DDASAC. Para tolerâncias inferiores àquele valor os métodos alternado corrector e simultâneo corrector, implementados pela DASPK, são mais eficientes, sendo mesmo os únicos que permitem o cálculo das sensibilidades com tolerâncias inferiores a  $10^{-7}$ .

## C.4 Conclusões

O estudo feito sobre a eficiência do cálculo das sensibilidades permitiu-nos obter resultados importantes que permitem uma redução significativa do esforço computacional necessário para a resolução do problema de controlo do processo TE. A análise feita relativamente às tolerâncias com que as sensibilidades são calculadas permitiu-nos concluir que a utilização de tolerâncias próximas das impostas ao cálculo das variá-

veis de estado resulta na necessidade de calcular estas com erros muito inferiores aos pretendidos, o que implica um esforço computacional adicional. Os resultados obtidos mostram que a utilização de uma tolerância de  $10^{-5}$  no cálculo das sensibilidades é um bom compromisso entre a rapidez e a precisão com que este cálculo é feito.

O estudo dos sistemas lineares envolvidos no cálculo numérico das variáveis de estado e das sensibilidades mostra que a matriz do Jacobiano do processo TE não é suficientemente esparsa ou estruturada para que qualquer uma destas características possa ser explorada. Todos os métodos testados neste sentido ou falharam ou produziram resultados, em termos de eficiência, piores do que os obtidos quando a matriz do Jacobiano é tratada como densa.

A comparação entre os diferentes métodos utilizados para o cálculo das sensibilidades mostra que o método alternado directo, implementado pela DDASAC, é o mais eficiente para a tolerância de  $10^{-5}$ . Dos resultados obtidos podemos concluir que o aproveitamento da aproximação da matriz do Jacobiano utilizada no cálculo das variáveis de estado para o cálculo das sensibilidades nos métodos alternado corrector e simultâneo corrector, implementados pela DASPK, não resulta num aumento da eficiência. Tal deve-se ao facto de este reaproveitamento resultar num maior número de passos do integrador, o que implica a necessidade de resolver um maior número de sistemas de equações lineares, relativamente ao método alternado directo. Daqui podemos concluir que, para o processo TE, o esforço computacional necessário à obtenção das sensibilidades é dominado pela factorização e solução de sistemas equações lineares e não pelo cálculo do Jacobiano, que no problema estudado é feito numericamente. Generalizando, podemos afirmar que a utilização dos métodos alternado corrector e simultâneo corrector apenas é vantajosa quando o cálculo do Jacobiano é computacionalmente dispendioso, quando comparado com o esforço necessário para factorizar e resolver sistemas de equações lineares. Na situação inversa, é vantajosa a utilização do método alternado directo uma vez que o cálculo do Jacobiano em cada passo do integrador permite uma redução significativa do número de passos necessários para a convergência.