



**THÈSE DE DOCTORAT**

**SPECIALITE : PHYSIQUE**

*Ecole Doctorale « Sciences et Technologies de l'Information des  
Télécommunications et des Systèmes »*

Présentée par :

**Ghizlane HAFIDI**

Sujet :

**Application de la commande prédictive non-linéaire  
à la commande de culture de bactéries *Escherichia coli***

Soutenue le 29 Septembre 2008 devant les membres du jury :

M Didier DUMUR

Directeur de thèse

Mme Françoise LAMNABHI-LAGARRIGUE

Présidente du Jury

M Mohammed M'SAAD

Rapporteur

Mme Isabelle QUEINNEC

Rapporteur

Mme Sihem TEBBANI

Co-directrice de thèse

M Alain VANDE WOUWER

Examineur



*A mes parents,  
A Imane,  
A ma belle-famille,  
A Omar,*



# Remerciements

Mes travaux de thèse se sont déroulés au Département d'Automatique de l'Ecole Supérieure d'Electricité (Supélec), dirigé par Monsieur Patrick Boucher, Professeur à Supélec. Je lui adresse ma reconnaissance pour m'avoir accueillie dans son département, encouragée à participer à des conférences internationales et d'avoir veillé sur les bonnes conditions de travail qui m'ont permis de mener à bien cette thèse.

Je remercie chaleureusement Mme Sihem Tebbani et Monsieur Didier Dumur, Professeurs à Supélec, qui ont encadré mes travaux de thèse avec professionnalisme, rigueur, patience, disponibilité et générosité. Je leur apporte ma plus sincère gratitude pour leur soutien constant, leurs encouragements et le temps précieux qu'ils m'ont accordé pendant ces trois années de thèse.

Cette thèse s'est déroulée en collaboration avec la faculté Polytechnique de Mons en Belgique où j'ai effectué plusieurs séjours. Je voudrai remercier le Professeur Alain Vande Wouwer de m'avoir accueillie au sein du Service d'Automatique de cette faculté et de m'avoir fait l'honneur d'être membre de mon jury. Je le remercie pour les longues discussions que nous avons eues ensemble et qui ont contribué à mener à bien mes travaux de thèse. Je tiens également à remercier Madame Anne-Lise Hantson, Professeur à la Faculté Polytechnique de Mons, qui m'a fait découvrir l'univers des biologistes. Je la remercie pour sa sympathie, sa gentillesse et sa disponibilité.

Je remercie Madame Françoise Lamnabhi-Lagarrigue, Directeur de recherches au CNRS, de m'avoir fait l'honneur de présider mon jury. Je tiens à rendre hommage à mes rapporteurs : Mme Isabelle Queinnec, HDR, Chargée de Recherches CNRS au LAAS-CNRS et Monsieur Mohammed M'Saad, Professeur à l'ENSICAen, pour le temps précieux qu'ils ont consacré à la lecture du présent mémoire.

Pendant ma thèse, j'ai encadré des élèves en travaux de laboratoire (TL) à Supélec sous la direction du Professeur Gilles Duc, que je remercie infiniment pour la confiance qu'il m'a accordée. Je remercie également Madame Houria Siguerdidjane, Professeur à Supélec, avec qui j'ai encadré plusieurs séances de TL, pour sa sympathie et sa gentillesse.

Je tiens à remercier tout particulièrement Madame Josiane Dartron, Secrétaire du Département d'Automatique, pour son soutien constant, ses encouragements, ses conseils, sa gentillesse et sa bonté. Je lui adresse mes amitiés les plus sincères.

Je n'oublierai pas de remercier les docteurs et doctorants des Départements d'Automatique de Supélec et de la Faculté Polytechnique de Mons pour tous les bons moments que nous avons partagés : Nicolas, Karim, Bastien, Cristina, Bilal, Mara, Giuliana, Lilian, Hichem, Ana-Maria, Florian, Henri, Olivier, Pedro, Sorin, Omar, Johan, Fransisca, Laurent... Un grand merci également à mes amis, Fadela, Kamal, Jamila, Claire, Maëlle, Marie-Noëlle, Yolande, Ilham, Emilie, Hala, Wissam, Rahma, Tarik, Samah et Zaatat.

Enfin, je remercie mes parents et ma sœur Imane pour leur amour, leurs encouragements, leur patience, leurs conseils et leur confiance indéfectible en moi. Je voudrai les remercier pour les valeurs qu'ils m'ont inculquées, telles que le sérieux et la persévérance, deux qualités essentielles pour mener à bien mes projets. J'associerai à ces remerciements toute ma famille, notamment mes grands-parents, mes futurs beaux-parents, Mima, Ali et Hind.

Pour finir, je voudrai ajouter que ce travail n'aurait pas pu être achevé sans l'amour, le soutien et les encouragements de mon futur mari, Omar. Je le remercie de m'avoir aidée à affronter les moments difficiles. Cette thèse représente pour moi l'aboutissement d'une première épreuve que nous avons traversée avec succès tous les deux. Je le remercie également pour l'intérêt particulier qu'il porte à ma carrière professionnelle. Je voudrai qu'il trouve ici tout l'amour que j'éprouve pour lui.

# Avant-propos

Le travail présenté dans cette thèse a donné lieu à la publication d'un certain nombre d'articles à l'occasion de différents congrès internationaux avec actes et sans actes, d'un article de revue et d'une communication nationale sans actes.

## Article de revue international avec comité de lecture

- S. Tebbani, D. Dumur, G. Hafidi, "Open-loop optimization and trajectory tracking of a fed-batch bioreactor", *Chemical Engineering and Processing: Process Intensification*, vol. 47-11, pp. 1933-1941, Octobre 2008.

## Conférences internationales avec actes

- G. Becerra-Celis, G. Hafidi, S. Tebbani, D. Dumur, A. Isambert, "Nonlinear predictive control for continuous microalgae cultivation process in a photobioreactor", *10<sup>th</sup> IEEE International Conference on Control, Automation, Robotics and Vision ICARCV 08*, Hanoi, Vietnam, décembre 2008.
- G. Hafidi, S. Tebbani, D. Dumur, A. Vande Wouwer, "Analyse de sensibilité pour l'identification d'un modèle paramétrique de culture de *E. coli*", *5<sup>ème</sup> Conférence Internationale Francophone d'Automatique (CIFA 08)*, CDRom Proceedings, Bucarest, Septembre 2008.
- G. Hafidi, S. Tebbani, D. Dumur, A. Vande Wouwer, "Nonlinear model Predictive Control applied to *E. coli* cultures", *17<sup>th</sup> International Federation of Automatic Control World Congress (IFAC WG 08)*, pp. 14570-14575, Seoul, Juillet 2008.
- G. Hafidi, S. Tebbani, D. Dumur, A. Vande Wouwer, "Robustness analysis of nonlinear predictive control of fed-batch *E. coli* cultures", *the 16<sup>th</sup> Mediterranean Conference on Control and Automation (MED 08)*, pp. 564-569, Ajaccio, Juin 2008.
- G. Hafidi, S. Tebbani, D. Dumur, A. Vande Wouwer, "Towards nonlinear model predictive control of fed-batch cultures of *E. coli*", *Multi-conference on Systems and Control MSC-CCA 2007*, pp. 194-199, Singapour, Octobre 2007. (Finaliste prix meilleur article étudiant)
- G. Hafidi, A.M. Vladu, S. Tebbani, D. Dumur, "Towards a generic model structure of bioreactors", *DYCOPS'07*, Vol. 2, pp. 255-260, Cancun, Juin 2007.
- S. Tebbani, D. Dumur, G. Hafidi, "Optimisation du fonctionnement d'un bioréacteur en mode fed-batch", *Systèmes d'Information, Modélisation, Optimisation et commande en génie des procédés SIMO06*, Toulouse, octobre 2006.

## Conférences internationales sans actes

- G. Hafidi, S. Tebbani, D. Dumur, A. Vande Wouwer, "Nonlinear predictive control of *E. coli* cultures", *NMPC-SOFAP workshop*, Loughborough, Avril 2007.

## Communications nationales sans actes

- G. Hafidi, S. Tebbani, D. Dumur, A. Vande Wouwer, "Commande prédictive non-linéaire de la bactérie *E. coli*", *GT Commande prédictive non-linéaire*, Juin 2007.



# Sommaire

1.	Introduction .....	13
1.1	Contexte .....	13
1.2	Motivations .....	14
1.3	Organisation de la thèse .....	15
2.	Modélisation du bioprocédé fed-batch de culture de bactéries <i>Escherichia coli</i> .....	19
2.1	Introduction .....	19
2.2	Description d'un bioprocédé .....	20
2.3	Modélisation mathématique d'un bioprocédé .....	21
2.3.1	Schéma réactionnel .....	22
2.3.2	Modèle dynamique général .....	22
2.3.3	Modélisation de la cinétique réactionnelle .....	23
2.3.4	Modélisation des transferts gazeux .....	25
2.4	Modélisation du bioprocédé fed-batch de culture de bactéries <i>Escherichia coli</i> .....	25
2.4.1	Schéma réactionnel .....	26
2.4.2	Modèle macroscopique .....	29
2.4.3	Cinétique des <i>E. coli</i> .....	31
2.5	Mise en œuvre d'un modèle simplifié du bioprocédé de culture de bactéries <i>E. coli</i> .....	34
2.6	Conclusions .....	35
3.	Vers une stratégie d'identification .....	39
3.1	Introduction .....	39
3.2	Identification basée sur le découplage .....	40
3.2.1	Identification des paramètres stœchiométriques .....	40
3.2.1.1	Méthode de découplage .....	40
3.2.1.2	Conditions de la « C-identifiabilité » .....	42
3.2.2	Identification des paramètres cinétiques .....	42
3.3	Identification basée sur l'analyse de sensibilité .....	43
3.3.1	Choix des paramètres à identifier .....	44
3.3.1.1	Analyse de sensibilité .....	44
3.3.1.2	Détermination des dépendances linéaires entre les paramètres .....	45
3.3.2	Procédure d'identification .....	46
3.4	Validation de l'identification .....	46
3.4.1	Intervalles de confiance .....	46
3.4.2	Erreur quadratique .....	46
3.5	Application de la stratégie basée sur le découplage au bioprocédé <i>E. coli</i> .....	47
3.5.1	Découplage de la stœchiométrie et de la cinétique du bioprocédé .....	47
3.5.2	Procédure d'identification paramétrique .....	50
3.6	Application de la stratégie basée sur l'analyse de sensibilité au bioprocédé <i>E. coli</i> .....	51

3.6.1	Détermination des paramètres à identifier .....	51
3.6.1.1	Fonctions de sensibilité du modèle du bioprocédé <i>E. coli</i> .....	52
3.6.1.2	Dépendances linéaires entre les paramètres.....	59
3.6.2	Procédure d'identification paramétrique.....	59
3.7	Résultats et validation en simulation.....	60
3.7.1	Stratégie basée sur le découplage.....	61
3.7.2	Stratégie basée sur l'analyse de sensibilité .....	62
3.7.3	Stratégie d'identification simultanée.....	62
3.7.4	Validation des modèles identifiés .....	63
3.8	Conclusions .....	64
4.	Détermination d'un profil optimal d'alimentation.....	65
4.1	Introduction .....	65
4.2	Formulation du problème .....	66
4.3	Résolution par l'approche directe.....	67
4.3.1	Approche séquentielle.....	67
4.3.2	Approche simultanée.....	67
4.4	Résolution par l'approche indirecte.....	68
4.4.1	Formulation du problème d'optimisation.....	68
4.4.2	Résolution.....	69
4.4.2.1	Méthode de tir .....	70
4.4.2.2	Méthode du gradient .....	70
4.4.2.3	Paramétrisation des états et des états adjoints.....	71
4.5	Récapitulatif des méthodes de résolution .....	72
4.6	Principe du Maximum de Pontryagin.....	73
4.7	Caractérisation analytique de la solution optimale.....	75
4.7.1	Caractérisation de la commande non singulière.....	75
4.7.2	Caractérisation de la commande singulière.....	76
4.7.2.1	Calcul analytique de la commande singulière .....	76
4.7.2.2	Détermination des intervalles singuliers.....	76
4.8	Application à la détermination d'un profil optimal de commande du bioprocédé <i>E. coli</i> .....	77
4.8.1	Mise en œuvre d'un modèle simplifié du bioprocédé <i>E. coli</i> .....	78
4.8.2	Définition du problème d'optimisation.....	80
4.8.3	Expressions analytiques des solutions optimales.....	81
4.8.3.1	Expression analytique de la commande non singulière $u_{path}$ .....	82
4.8.3.2	Expression analytique de la commande singulière $u_{sing}$ .....	83
4.8.4	Récapitulatif.....	84
4.9	Résultats numériques.....	85
4.10	Conclusions.....	88
5.	Mise en œuvre d'une stratégie de commande prédictive non-linéaire.....	91
5.1	Introduction .....	91
5.2	Commande prédictive.....	93
5.2.1	Principe de la commande prédictive .....	93
5.2.2	Bref historique de la commande prédictive.....	95
5.2.3	Commande prédictive non-linéaire .....	96
5.2.3.1	Définition de la fonction coût .....	96
5.2.3.2	Formulation du problème d'optimisation .....	97
5.3	Etude de la commandabilité du bioprocédé <i>E. coli</i> .....	98
5.3.1	Conditions de commandabilité.....	98
5.3.2	Commandabilité du bioprocédé <i>E. coli</i> .....	98

5.4	Application de la commande prédictive non-linéaire au bioprocédé <i>E. coli</i> .....	101
5.4.1	Objectif de commande .....	101
5.4.2	Formulation de la commande prédictive non-linéaire appliquée au bioprocédé <i>E. coli</i> .....	102
5.4.2.1	Définition du critère.....	102
5.4.2.2	Prise en compte d'une erreur de modèle.....	103
5.4.3	Résolution du problème d'optimisation .....	104
5.5	Analyse des performances au nominal .....	107
5.5.1	Régulation de la concentration en acétate à $A_{set} = 0$ g/kg .....	107
5.5.2	Régulation de la concentration en acétate à $A_{set} = 0,5$ g/kg.....	109
5.5.2.1	Détermination du profil d'alimentation sous-optimal associé.....	109
5.5.2.2	Résultats de simulation .....	111
5.5.3	Régulation de la concentration en acétate avec changement de consigne .....	116
5.5.4	Influence des paramètres de commande.....	121
5.6	Comparaisons avec d'autres commandes .....	124
5.6.1	Comparaison avec la commande adaptative linéarisante .....	124
5.6.1.1	Commande adaptative linéarisante appliquée au bioprocédé <i>E. coli</i> .....	125
5.6.1.2	Résultats de simulation .....	126
5.6.2	Comparaison avec la commande par modèle générique (GMC) .....	129
5.6.2.1	Commande par modèle générique appliquée au bioprocédé <i>E. coli</i> .....	130
5.6.2.2	Résultats de simulation .....	130
5.7	Conclusions .....	133
6.	Etude de la robustesse de la loi de commande prédictive non-linéaire.....	135
6.1	Introduction .....	135
6.2	Robustesse vis-à-vis d'erreurs de modèle .....	136
6.2.1	Analyse d'un scénario particulier.....	136
6.2.1.1	Prise en compte de la différence objet-modèle.....	136
6.2.1.2	Utilité de la différence objet-modèle .....	138
6.2.2	Analyse statistique.....	140
6.3	Analyse de la robustesse vis-à-vis du bruit de mesure .....	141
6.4	Comparaison avec d'autres commandes.....	142
6.4.1	Comparaison avec la commande adaptative linéarisante.....	142
6.4.2	Comparaison avec la commande par modèle générique (GMC) .....	144
6.5	Conclusions .....	145
7.	Conclusions - Perspectives.....	147
7.1	Bilan du travail effectué .....	147
7.2	Apports scientifiques et originalité du travail.....	149
7.3	Perspectives .....	150
8.	Références bibliographiques .....	153
9.	Annexes.....	161
9.1	Annexe 1 : Analyse de sensibilité en régime oxydatif .....	161
9.1.1	Fonctionnement en sous-régime 1.....	161
9.1.1.1	Bioréacteur en mode batch.....	161
9.1.1.2	Bioréacteur en mode fed-batch .....	164
9.1.2	Fonctionnement en sous-régime 2.....	166
9.1.2.1	Bioréacteur en mode batch.....	167
9.1.2.2	Bioréacteur en mode fed-batch .....	169
9.1.3	Conclusions .....	171
9.2	Annexe 2 : Expression détaillée de la matrice $M$ .....	172
9.3	Annexe 3 : Expression de $S_{crit}$ en fonction de $A_{set}$ .....	175

9.4	Annexe 4 : Montage expérimental.....	176
9.4.1	Présentation du bioréacteur .....	176
9.4.2	Dispositifs de mesures.....	177
9.4.2.1	Mesure de la concentration en biomasse .....	177
9.4.2.2	Mesure de la concentration en glucose .....	177
9.4.2.3	Mesure de la concentration en acétate .....	178
9.4.2.4	Mesures et régulations du pH et de l'oxygène dissous.....	178
9.4.2.5	Mesure du poids.....	178
9.4.3	Préparation d'une préculture .....	179
9.4.4	Réalisation d'une expérience « culture » .....	181

# 1. Introduction

## 1.1 Contexte

Les travaux effectués au cours de cette thèse se situent à la croisée d'un domaine théorique très important, l'Automatique et la commande des systèmes plus particulièrement, et d'un secteur applicatif de tout premier plan, le secteur des biotechnologies.

Les biotechnologies, selon l'Organisation de Coopération et de Développement Economiques (OCDE), sont définies comme « l'application de la science et de la technologie aux organismes vivants et à d'autres matériaux vivants ou non vivants, pour la production de savoir, biens et services ». Elles sont utilisées par l'homme depuis des siècles. Par exemple, les micro-organismes étaient utilisés pour la fabrication des produits alimentaires tels que le pain, le fromage et les boissons alcoolisées.

De nos jours, les biotechnologies permettent à un industriel d'améliorer ses procédés de fabrication, en modernisant ou en remplaçant des techniques plus anciennes. Elles permettent également de créer des produits inconcevables auparavant. L'introduction des biotechnologies implique donc une baisse significative des coûts de production et une augmentation de la qualité des produits. Tous ces aspects constituent un véritable enjeu économique.

Les biotechnologies interviennent dans plusieurs domaines, notamment les domaines de l'agroalimentaire et de la protection de l'environnement. Dans les brasseries par exemple, les biotechnologies se sont substituées aux techniques artisanales. Le contrôle de la fabrication et la sélection des levures assurant la fermentation permettent désormais d'obtenir une bière de qualité et de goût constants. Les enzymes introduites dans les lessives permettent de laver le linge à basse température, sans ajouter de phosphates, diminuant ainsi la consommation d'énergie et limitant la pollution. Les micro-organismes sont également utilisés pour dégrader la pollution pour l'épuration des eaux.

Les biotechnologies interviennent aussi dans le domaine de la santé (pharmaceutique et médical). Il y a quelques années, la fabrication des médicaments, des vaccins et des traitements était réalisée uniquement grâce à la chimie. Maintenant, les biotechnologies commencent à prendre une part importante dans la fabrication des médicaments et vaccins, par exemple, la production d'insuline et d'hormones de croissance. Toutefois, le rôle le plus important qu'elles occupent dans le domaine de la santé est la découverte de nouveaux traitements qui sont trop coûteux ou impossibles à obtenir à partir d'une synthèse chimique.

Elles sont également utilisées pour résoudre des problèmes de détection et de traçabilité. Le génie génétique est notamment exploité dans le cadre de la répression des fraudes, par exemple, pour s'assurer de l'absence d'OGM dans un aliment, ou la recherche de paternité en comparant les génomes ou encore par la police scientifique pour la comparaison des ADN.

Les applications des biotechnologies sont larges et variées mais elles constituent toutes un vrai défi économique. C'est ainsi que le domaine des biotechnologies est devenu l'un des domaines les plus étudiés par la communauté automatique.

De l'autre côté, des lois de commande de plus en plus sophistiquées ont vu le jour, permettant d'atteindre des performances de plus en plus poussées, certes parfois au prix d'une complexité

accrue notamment en termes de réglage. L'essor des moyens informatiques a joué pour beaucoup dans ce développement théorique. Les stratégies transférables à l'industrie sont donc de plus en plus nombreuses, en particulier vers des secteurs pour lesquels le plus souvent aucune stratégie avancée n'est implantée. C'est en particulier le cas des biotechnologies, où l'on peut trouver des régulations très simples pour le pH ou la pression en oxygène, mais rien en termes de commande avancée pour la régulation de certaines concentrations par exemple. Parmi toutes ces stratégies, la commande prédictive est une technique très implantée dans le milieu industriel, simple et performante, adaptée notamment aux systèmes complexes non-linéaires tels que ceux rencontrés dans le domaine des biotechnologies. Dans la littérature, depuis une quinzaine d'années, de nombreuses publications ont trait à l'utilisation de la commande prédictive linéaire et non-linéaire à des bioréacteurs en modes batch et fed-batch. Parmi ces nombreuses publications, on trouvera plus particulièrement en référence celles ayant trait spécifiquement à la culture des bactéries *E. coli*. L'état de l'art sur les *E. coli*, assez restreint, propose souvent des solutions numériques au problème de commande, notre travail se démarque par une contribution analytique au niveau de la mise en œuvre de la commande (trajectoire optimale en particulier).

## 1.2 Motivations

A partir du contexte général exposé précédemment, le domaine d'application des biotechnologies qui nous intéresse ici est la santé, au travers de la culture des bactéries *Escherichia coli*. Ces bactéries sont en effet très utilisées par les généticiens pour fabriquer des médicaments ou des vaccins car elles ont l'avantage d'avoir une croissance relativement rapide par rapport aux autres micro-organismes, elles se divisent toutes les 20 minutes. Ces bactéries sont utilisées notamment pour la production d'insuline, de vaccins pédiatriques contre la méningite et la typhoïde et de quelques anti-virus. Il est donc important de produire des quantités de médicaments et de vaccins à l'échelle industrielle.

La croissance des bactéries *Escherichia coli* est réalisée à l'intérieur d'un bioréacteur pendant une durée donnée nommée « batch ». La biomasse constituée de la bactérie est extraite à la fin du batch et utilisée pour la fabrication des médicaments ou vaccins selon la souche de bactérie utilisée initialement. Il est donc important de produire une grande quantité de biomasse durant un batch. On comprend alors aisément la motivation de notre travail, visant à augmenter les performances de fonctionnement des bioprocédés *E. coli*, en introduisant des lois de commande performantes afin de répondre aux besoins des industriels.

Comme mentionné précédemment, la commande des bioprocédés se limite très souvent à des régulations permettant d'avoir des conditions environnantes favorables (température, pH, oxygène...). Or la régulation d'un des constituants du système constitue une étape primordiale pour augmenter les performances des bioprocédés.

L'objectif de cette thèse consiste alors à appliquer une technique de commande avancée, à savoir la commande prédictive non-linéaire au bioprocédé fed-batch de culture de bactéries *E. coli* afin de maximiser au mieux la croissance de la biomasse.

## 1.3 Organisation de la thèse

La structure de cette thèse suit graduellement les étapes nécessaires à l'élaboration de la loi de commande prédictive non-linéaire pour la culture fed-batch de bactéries *E. coli*. Le travail effectué a pour objectif d'apporter une réponse la plus complète possible à la problématique de maximisation de la croissance des bactéries. Cela passe bien sûr par l'élaboration d'une stratégie de commande avancée. Mais cette stratégie nécessitant l'utilisation d'un modèle pour sa mise en œuvre fait que l'on se doit avant tout de développer une structure d'identification des paramètres du modèle. Ce sera l'objet d'un chapitre spécifique, même si cette thèse ne se veut pas une « thèse d'identification ». De même, la stratégie de commande fait appel au suivi d'un profil d'alimentation de référence, élaboré à l'issue d'une recherche du profil optimal. Là encore, un chapitre spécifique est dédié à cette recherche. Pour chacun de ces deux aspects (identification et profil optimal), des travaux restent encore à accomplir, ce seront des perspectives dans la continuité de la thèse.

Globalement donc, ce mémoire propose tous les outils nécessaires permettant une production maximale de bactéries *E. coli* au sein d'un bioréacteur opérant en mode fed-batch. En plus de ce court chapitre d'introduction et d'une partie dédiée aux conclusions et perspectives, ce mémoire est divisé en cinq chapitres dont le détail est donné ci-dessous.

### *Chapitre 2*

Le deuxième chapitre a pour objectif l'élaboration d'un modèle mathématique macroscopique décrivant le fonctionnement du bioprocédé. Pour cela, des notions liées aux bioprocédés sont tout d'abord présentées pour familiariser le lecteur avec ce domaine. Puis, le modèle macroscopique du bioprocédé est déduit du schéma réactionnel, il sera utilisé pour les validations en simulation comme modèle complet de simulation. Par ailleurs, ce modèle est également simplifié pour aboutir à un modèle de commande repris dans toutes les études théoriques présentées dans les chapitres ultérieurs, à savoir l'identification des paramètres du modèle, la recherche d'un profil optimal et la synthèse de la loi de commande prédictive non-linéaire.

### *Chapitre 3*

Ce chapitre est dédié à l'identification des paramètres du modèle de commande. La mise en œuvre d'une procédure d'identification s'avère de première importance car la synthèse de la loi de commande prédictive nécessite la connaissance d'un modèle, même imparfait, du système à piloter. Deux stratégies d'identification sont proposées dans ce chapitre. La première reprend une démarche connue dans la littérature et repose sur l'idée d'identifier séparément les paramètres stœchiométriques et les paramètres de la cinétique (découpler la stœchiométrie et la cinétique). La deuxième, originale dans ces travaux, se base sur l'analyse de sensibilité des sorties du modèle vis-à-vis de ses paramètres. Les procédures d'identification sont décrites ainsi que les outils de validation associés. Les résultats numériques obtenus à partir de ces deux stratégies sont comparés aux résultats d'une approche classique consistant à identifier l'ensemble des paramètres de façon globale.

#### **Chapitre 4**

Le quatrième chapitre a pour objectif de déterminer un profil d'alimentation optimal du bioréacteur fed-batch de culture de bactéries *E. coli* en résolvant un problème d'optimisation pour maximiser la quantité de biomasse. Ce profil sera utilisé par la suite par le critère considéré pour la synthèse de la loi de commande prédictive. L'application du Principe du Maximum de Pontryagin permet de caractériser toutes les solutions optimales possibles. Cependant, compte tenu de la complexité du problème et de la nature hybride du procédé (différents régimes de fonctionnement), la sélection de la solution optimale parmi toutes les expressions possibles pour maximiser la quantité de biomasse est effectuée par le biais de résultats de simulation. Le résultat obtenu est corroboré par une analyse biologique du bioprocédé, montrant également que la maximisation de la quantité de biomasse se traduit par la régulation de la concentration en acétate à une valeur très proche de zéro. Le profil optimal d'alimentation déterminé dans ce chapitre servira donc comme trajectoire de référence pour la commande prédictive non-linéaire élaborée au chapitre suivant.

#### **Chapitre 5**

Ce chapitre présente la mise en œuvre d'une stratégie de commande prédictive non-linéaire pilotant le bioprocédé *E. coli* et permettant la régulation de la concentration en acétate à une valeur aussi proche que possible de zéro, tout en forçant le débit d'alimentation à suivre un profil de référence donné. Le profil d'alimentation optimal déterminé au chapitre 4 est utilisé dans un premier temps comme trajectoire de référence pour une régulation en acétate à une valeur nulle. Mais, compte-tenu de la sensibilité des capteurs, ce cas de figure s'avère peu réaliste en pratique, nécessitant alors la détermination d'un profil d'alimentation sous-optimal tel que le bioprocédé fonctionne à la frontière des deux régimes (régime oxydatif et régime oxydo-fermentatif). Ce profil est calculé hors ligne pour un point d'équilibre correspondant à des concentrations en acétate et en substrat données.

La synthèse de la commande proposée ici, incluant une trajectoire de référence en entrée et en sortie, est réalisée en plusieurs étapes. Tout d'abord et afin d'éviter les problèmes liés à la discrétisation du modèle et l'augmentation de la charge de calcul en temps réel, le problème d'optimisation classique de la commande prédictive non-linéaire est transformé en un problème de programmation non-linéaire, résolu par des techniques de CVP (Control Vector Parametrization). Puis un changement de variable permet de se ramener à un problème d'optimisation non contraint. Enfin, la structure proposée prend en compte également la différence entre le modèle et le bioprocédé, conférant ainsi à la structure asservie un certain degré de robustesse vis-à-vis d'erreurs paramétriques.

La loi de commande proposée est ensuite validée dans le cadre de plusieurs scénarii. Les performances obtenues à partir de cette loi de commande sont au final comparées à celles obtenues avec deux stratégies rencontrées dans le domaine des bioprocédés, la commande adaptative linéarisante et la commande par modèle générique.

#### **Chapitre 6**

Ce chapitre propose une étude de la robustesse de la loi de commande prédictive non-linéaire élaborée au chapitre précédent. Cette analyse est principalement basée sur des résultats obtenus pour plusieurs scénarii permettant de valider la robustesse de la loi de commande

prédictive vis-à-vis de variations de paramètres et vis-à-vis de bruits de mesure. L'intérêt de la prise en compte de la différence entre le modèle et le bioprocédé est également établi. Une analyse statistique par une approche type Monte Carlo vient confirmer l'ensemble des conclusions. Enfin, les performances en robustesse de la loi de commande prédictive non-linéaire sont comparées à celles issues des deux stratégies présentées au chapitre précédent : commande adaptative linéarisante et commande par modèle générique.

### ***Conclusions – Perspectives***

Cette dernière partie dresse le bilan des développements présentés dans ce mémoire et proposent des perspectives dans la continuité des travaux effectués, à la fois au niveau théorique et au niveau plus appliqué, en particulier en ce qui concerne l'implantation de la stratégie prédictive sur un bioréacteur de laboratoire.

### ***Annexes***

Ce mémoire inclut quatre annexes, une première décrivant l'analyse de sensibilité effectuée en régime oxydatif, en complément de celle donnée dans le corps du mémoire pour le régime oxydo-fermentatif, une deuxième regroupant les expressions des termes de la matrice intervenant dans l'étude de la commande singulière lors de la recherche du profil optimal d'alimentation, une troisième détaillant le calcul des couples  $(A_{set} \ S_{crit})$  pour lesquelles les trajectoires de référence sont calculées. La dernière annexe donne un descriptif du dispositif expérimental disponible au laboratoire de biochimie de la Faculté Polytechnique de Mons en Belgique pour le test ultérieur de la loi de commande.



## 2. Modélisation du bioprocédé fed-batch de culture de bactéries *Escherichia coli*

### 2.1 Introduction

La modélisation mathématique des bioprocédés s'avère relativement complexe car, contrairement à d'autres domaines, il n'existe pas de lois clairement admises décrivant l'évolution des micro-organismes. Cependant, modéliser est indispensable pour l'amélioration des performances des bioprocédés, intervenant dans de nombreux secteurs économiques très importants (santé, agroalimentaire, biocarburants, traitement des eaux...). En effet, la modélisation mathématique est à la base du développement des algorithmes d'optimisation, de commande, de supervision et de diagnostic, permettant tous un gain en performances.

Ce travail de thèse s'intéresse à la commande du bioprocédé de culture de bactéries *Escherichia coli*, il est donc nécessaire de déterminer un modèle mathématique de ce bioprocédé. On trouve classiquement dans la littérature deux types de modèles :

- les modèles dits structurés se basent sur une description détaillée de toutes les réactions physico-chimiques entre les différents constituants intracellulaires des micro-organismes. Leurs schémas réactionnels font intervenir plusieurs réactions et plusieurs constituants et il en résulte des modèles très complexes,
- les modèles dits non-structurés se basent sur une description macroscopique des réactions entre les constituants extracellulaires. Dans ce cas, les schémas réactionnels et les modèles sont plus simples que les précédents. C'est pour cette raison que les automaticiens s'intéressent à ce type de modélisation. Elle leur permet, contrairement aux modèles structurés, de développer les différentes techniques de l'Automatique. Dans ce sens, [Bastin et Dochain, 1990] est une référence incontournable pour la modélisation des bioprocédés.

Enfin, du point de vue de l'automaticien, il convient aussi de distinguer entre modèle de simulation, par exemple un modèle non structuré décrit précédemment, permettant de simuler le comportement du bioprocédé le plus fidèlement possible ; et modèle de commande, issu du modèle de simulation après simplifications, devant permettre l'élaboration de la loi de commande. Il est clair par ailleurs que la structure de commande sera testée préalablement à toute implantation expérimentale sur le modèle de simulation et devra posséder de bonnes propriétés de robustesse pour induire des performances satisfaisantes sur un modèle complet, sachant qu'elle résulte d'une synthèse sur un modèle simplifié.

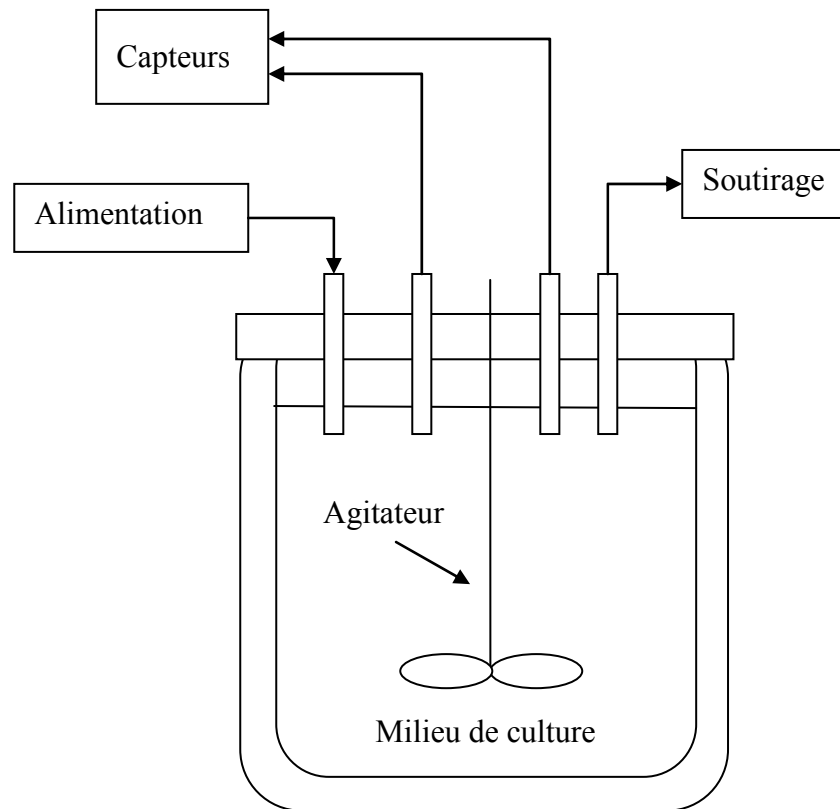
Les paragraphes ci-dessous visent à élaborer, dans le cas de la modélisation du bioprocédé fed-batch de culture de bactéries *E. coli*, ces modèles non structurés de simulation et de commande. Une partie préalable rappelle la description d'un bioprocédé, les différents modes de fonctionnement ainsi que la structure générale d'un modèle dynamique associé à un bioprocédé, se basant sur un schéma réactionnel et tenant compte de la cinétique réactionnelle et des transferts gazeux.

## 2.2 Description d'un bioprocédé

Un bioprocédé est une structure qui consiste à faire croître une population de micro-organismes (bactéries, levures, moisissures...) en présence de certains éléments nutritifs (nutriments) dans un environnement (température, pH, agitation, aération ...) favorable. La croissance des micro-organismes à l'intérieur d'un bioprocédé (milieu de culture) est le résultat de l'activité métabolique des cellules.

On s'intéresse aux bioprocédés infiniment mélangés uniquement. La Figure 2.1 montre un schéma représentatif classique d'un bioprocédé. Il est composé :

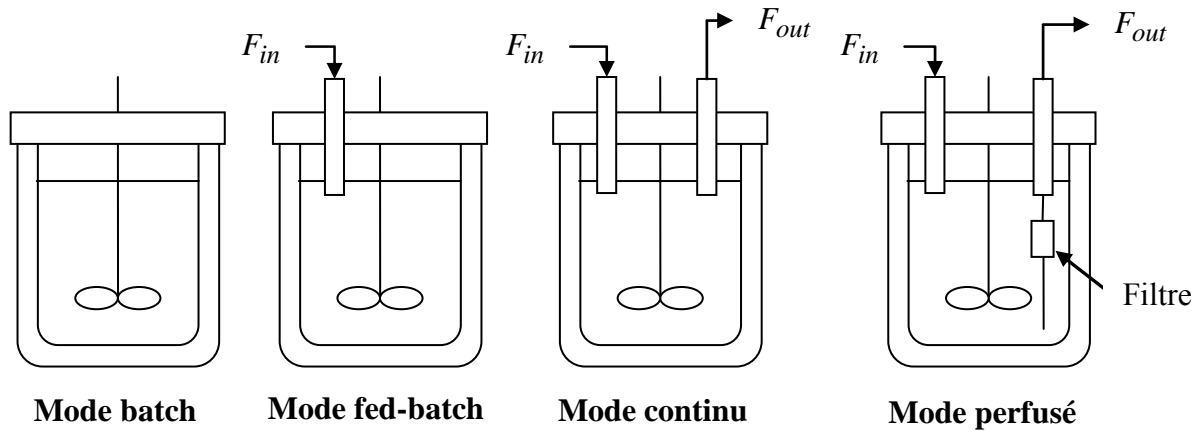
- d'une double paroi dans laquelle circule de l'eau permettant la régulation de la température à l'intérieur du bioprocédé,
- d'un agitateur permettant l'homogénéité du milieu de culture,
- de plusieurs connecteurs permettant l'ajout du milieu d'alimentation, le soutirage du milieu de culture et la liaison avec les différents capteurs réalisant les mesures en ligne du pH, de la température et/ou des concentrations des divers constituants.



**Figure 2.1.** Schéma représentatif d'un bioprocédé.

A ce jour, la plupart des bioprocédés sont équipés de capteurs permettant les régulations du pH, de la température et de la concentration en oxygène dissous, pour permettre au bioprocédé de fonctionner dans des conditions favorables. Le pH est régulé en ajoutant une solution basique ou acide au milieu de culture. La température est régulée à l'aide de l'eau circulant dans la double paroi, enfin la concentration en oxygène dissous est régulée en agissant sur la vitesse d'agitation et/ou le débit d'air injecté dans le milieu de culture.

On rencontre pour un bioprocédé quatre modes de fonctionnement. La différence provient principalement des types de débit d'alimentation  $F_{in}$  et de soutirage du milieu de culture  $F_{out}$ . La Figure 2.2 représente un schéma de ces quatre modes de fonctionnement.



**Figure 2.2.** Différents modes de fonctionnement d'un bioprocédé.

- mode batch (ou discontinu) : Les éléments nutritifs et la biomasse sont introduits au démarrage de la culture, et les débits d'alimentation et de soutirage du milieu de culture sont nuls (aux prélèvements d'échantillons près). Le volume du milieu de culture est donc approximativement constant. En effet, en pratique, une légère variation du volume est due à la régulation du pH et à l'évaporation (échange entre la phase liquide et la phase gazeuse). La biomasse et le produit ne sont récupérés qu'à la fin de la culture.
- Mode fed-batch (ou semi-continu) : Les éléments nutritifs sont introduits au fur et à mesure des besoins des micro-organismes. Le débit d'alimentation est non nul et le débit de soutirage est nul. Ainsi, le volume du milieu de culture augmente et la biomasse et le produit sont récupérés à la fin de la culture. Pour une meilleure exploitation de ce mode, le débit d'alimentation doit être calculé à l'aide d'un régulateur.
- Mode continu (ou chemostat) : Les éléments nutritifs sont introduits de manière continue et le milieu de culture est soutiré continuellement, de façon telle que le volume du milieu de culture reste constant. Les débits d'alimentation et de soutirage sont donc non nuls. Ce mode est par exemple employé dans le cas du traitement des eaux.
- Mode perfusé : Comme pour le mode continu, les débits d'alimentation et de soutirage sont non nuls. La principale différence est la présence du filtre permettant d'évacuer le milieu de culture usé et de garder la biomasse à l'intérieur du bioprocédé. Ce mode est utilisé pour les cultures de cellules animales.

## 2.3 Modélisation mathématique d'un bioprocédé

[Bastin et Dochain, 1990] est une référence incontournable pour la modélisation mathématique des bioprocédés. Les auteurs y introduisent en particulier la notion de schéma

réactionnel et proposent un modèle mathématique générique des bioprocédés dit « modèle dynamique général ».

Ce paragraphe récapitule les principaux aspects liés à la modélisation mathématique, à savoir le schéma réactionnel, le modèle dynamique général proposé par Bastin et Dochain, la modélisation de la cinétique réactionnelle et des transferts gazeux utilisée dans le domaine des bioprocédés. Les notations adoptées sont celles mises en œuvre par Bastin et Dochain.

### 2.3.1 Schéma réactionnel

La notion de schéma réactionnel d'un bioprocédé, introduite dans [Bastin et Dochain, 1990], est une représentation qualitative des relations entre les différents constituants macroscopiques (réactifs, produits et/ou catalyseurs).

Un bioprocédé est un ensemble de  $n_r$  réactions impliquant  $n$  constituants. Le schéma réactionnel générique s'écrit comme suit :



où :

- $\xi_i$  et  $\xi_j$  sont respectivement les concentrations du  $i^{\text{ème}}$  constituant pour les réactifs et du  $j^{\text{ème}}$  pour les produits,
- $k_{il}$  et  $k_{jl}$  représentent les coefficients pseudo-stœchiométriques ; ils sont positifs,
- $r_l$  est la vitesse de réaction  $l$ ,
- $R_l$  et  $P_l$  sont respectivement les ensembles de constituants  $\xi_i$  et  $\xi_j$  dans la réaction  $l$ ,
- $n_r$  est le nombre de réactions.

L'utilisation du schéma réactionnel (2.1) nécessite les précautions suivantes :

- il ne représente pas les relations stœchiométriques entre les constituants comme c'est le cas pour les réactions chimiques. C'est la raison pour laquelle les coefficients sont dits « pseudo-stœchiométriques »,
- la notation  $\xi_i$  regroupe quatre types de constituants, micro-organismes, enzymes, substrats extérieurs et substrats produits par une ou plusieurs réactions et utilisés par une ou d'autres réactions,
- le schéma réactionnel (2.1) ne représente pas une description exhaustive du bioprocédé. Les substrats non limitants et les sous-produits sont souvent omis quand ils n'interviennent pas dans d'autres réactions.

### 2.3.2 Modèle dynamique général

Le modèle dynamique introduit par Bastin et Dochain se base sur le schéma réactionnel présenté précédemment. A partir du schéma réactionnel (2.1), l'équation différentielle décrivant la dynamique de la concentration  $\xi_i$  s'écrit comme suit :

$$\frac{d\xi_i}{dt} = \sum_{j \sim i} (\pm) k_{ij} r_j - D\xi_i - Q_i + F_i \quad (2.2)$$

où :

- la notation  $j \sim i$  signifie que la sommation est effectuée sur les réactions d'indice  $j$  qui impliquent le constituant d'indice  $i$ ,
- les coefficients pseudo-stœchiométriques  $k_{ij}$  sont précédés d'un signe négatif quand  $\xi_i$  représente un réactif dans la réaction  $j$  et d'un signe positif quand il représente un produit dans la réaction  $j$ ,
- $D$  représente le taux de dilution,
- $Q_i$  est le débit massique volumique de sortie du constituant  $\xi_i$  sous forme gazeuse,
- $F_i$  est le débit massique volumique d'entrée du constituant  $\xi_i$  sous forme liquide ou gazeuse si  $\xi_i$  est un substrat externe. Sinon  $F_i = 0$ .

Introduisons les notations matricielles suivantes :

$$\xi^T = [\xi_1, \xi_2, \dots, \xi_n], \quad r^T = [r_1, r_2, \dots, r_{n_r}], \quad Q^T = [Q_1, Q_2, \dots, Q_n], \quad F^T = [F_1, F_2, \dots, F_n],$$

$$K = [K_{ij}] \text{ matrice } n \times n_r \text{ avec } K_{ij} = (\pm)k_{ij} \text{ si } j \sim i, \text{ sinon, } K_{ij} = 0.$$

A l'aide de ces notations, le modèle dynamique général [Bastin et Dochain, 1990] s'écrit sous la forme :

$$\frac{d\xi}{dt} = Kr(\xi, t) - D\xi - Q(\xi) + F \quad (2.3)$$

où :

- $\xi$  est le vecteur des variables d'état,
- $K$  est la matrice des coefficients pseudo-stœchiométriques,
- $r$  est le vecteur des vitesses de réaction dépendant du temps et des variables d'état,
- le terme  $Kr(\xi, t)$  décrit la cinétique des réactions biochimiques considérées dans le schéma réactionnel (2.1),
- Le terme  $-D\xi + F - Q$  décrit la dynamique associée aux phénomènes de transport des constituants considérés.

### 2.3.3 Modélisation de la cinétique réactionnelle

Les fonctions cinétiques décrivent en général les phénomènes d'activation, de limitation et d'inhibition. Elles se présentent couramment sous la forme de fonctions rationnelles des concentrations en constituants  $\xi_i$  et font intervenir un ensemble de paramètres cinétiques. Elles traduisent l'influence des différents constituants sur les vitesses de réaction.

Les phénomènes les plus décrits dans la littérature sont l'activation et la limitation ou l'inhibition par un substrat ( $S$ ), l'activation par la biomasse ( $X$ ) et l'inhibition par un produit métabolique. Quelques lois décrivant la cinétique réactionnelle ont été déterminées par les biologistes selon l'effet sur la croissance. Nous citerons ici les lois les plus classiques de Monod, Haldane, Contois et Herbert.

La loi de Monod a été introduite dès le début du siècle par Michaëlis-Menten pour décrire une réaction enzymatique ; elle traduit le phénomène de limitation de la croissance par manque de substrat  $S$  sans tenir compte de l'effet inhibiteur. Elle est fréquemment utilisée et s'écrit sous la forme suivante [Monod, 1942] :

$$\mu(S) = \mu_{\max} \frac{S}{k_s + S} \quad (2.4)$$

où :

- $\mu_{\max}$  est le taux de croissance maximal,
- $k_s$  est une constante de saturation.

La loi de Haldane a été introduite dans le cas des réactions enzymatiques et reprise par Andrews dans le cas des réactions biologiques [Andrews, 1968]. Elle décrit l'effet inhibiteur de croissance par un substrat  $S$  et s'écrit sous la forme suivante :

$$\mu(S) = \mu_{hal} \frac{S}{k_s + S + \frac{S^2}{k_i}} \quad (2.5)$$

$$\mu_{hal} = \mu_{\max} (1 + 2\sqrt{k_s / k_i})$$

où  $k_i$  est une constante d'inhibition.

Le taux de croissance maximal  $\mu_{\max}$  se produit pour une concentration  $S^* = \sqrt{k_s k_i}$ . Si l'effet inhibiteur de  $S$  est négligeable ( $k_i \gg k_s$ ), l'équation (2.5) se réduit à (2.4).

La loi de Contois, quant à elle, tient compte du ralentissement de croissance généralement observé pour les grandes concentrations en biomasse  $X$ . Elle se présente sous la forme :

$$\mu(X, S) = \mu_{\max} \frac{S}{k_c X + S} \quad (2.6)$$

où  $\mu_{\max}$  et  $k_c$  sont des constantes.

Outre l'activation et la limitation par  $S$ , cette loi prend en compte un éventuel effet inhibiteur de  $X$ .

La loi de Herbert tient compte de la maintenance et la mortalité des cellules, par exemple dans le cas de culture de cellules animales. Elle se présente sous la forme :

$$\mu(S) = \mu_{\max} \frac{S}{k_s + S} - m \quad (2.7)$$

où  $m$  est un coefficient de maintenance et/ou de mortalité.

### 2.3.4 Modélisation des transferts gazeux

La modélisation des transferts gazeux est connue sous le nom de loi de Henry, formulée en 1803 par William Henry :

*A température constante et à saturation, la quantité de gaz dissous dans un liquide est proportionnelle à la pression partielle qu'exerce ce gaz en ce point.*

Dans le cas des bioprocédés, les transferts gazeux concernent essentiellement l'oxygène et le dioxyde de carbone (éventuellement d'autres gaz spécifiques, par exemple le méthane).

Le taux de transfert d'oxygène de la phase gazeuse à la phase liquide est noté *OTR* (*Oxygen Transfer Rate*) et s'écrit :

$$OTR = (k_L a)^{O_2} (O_{sat} - O) \quad (2.8)$$

où :

- *O* représente la concentration en oxygène dissous,
- $k_L a$  représente le coefficient de transfert de masse de la phase liquide,
- *a* représente la surface de l'interface liquide-gaz par unité de volume,
- $O_{sat}$  est la constante de saturation d'oxygène dans le milieu de culture.

Le taux de transfert du dioxyde de carbone de la phase liquide à la phase gazeuse est noté *CTR* (*Carbon dioxide Transfer Rate*) et s'écrit sous la forme :

$$CTR = (k_L a)^{CO_2} (C - C_{sat}) \quad (2.9)$$

où :

- *C* représente la concentration en dioxyde de carbone dissous,
- $C_{sat}$  est la constante de saturation du dioxyde de carbone dans le milieu de culture.

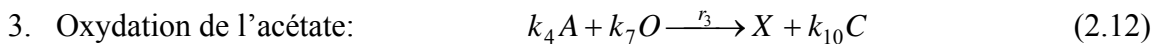
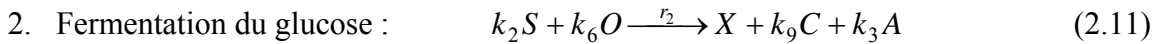
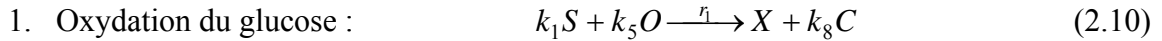
Les paramètres  $k_L a$ ,  $O_{sat}$  et  $C_{sat}$  dépendent fortement des conditions opératoires (agitation, débit de gaz, ajout d'agent anti-mousse,...), du milieu de culture (sels minéraux en particulier) et de l'état du milieu de culture (viscosité, température, pression, ...).

## 2.4 Modélisation du bioprocédé fed-batch de culture de bactéries *Escherichia coli*

Tous les concepts définis lors des paragraphes précédents vont maintenant être développés dans le cadre de nos travaux, à savoir le bioprocédé fed-batch de culture de bactéries *E. coli*. Une première étape permettra de définir le schéma réactionnel propre à ce type de bactéries, détaillera la théorie du « goulot d'étranglement ». Ce schéma réactionnel sera ensuite utilisé pour l'établissement du modèle macroscopique associé aux *E. coli*. Les cinétiques réactionnelles seront enfin décrites.

## 2.4.1 Schéma réactionnel

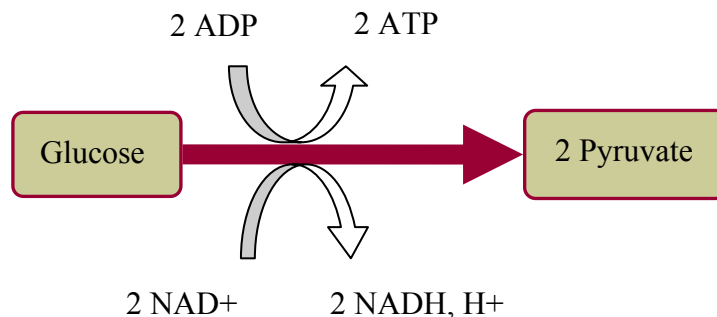
Le schéma réactionnel décrivant le métabolisme<sup>1</sup> des bactéries *E. coli* est constitué de trois voies métaboliques détaillées ci-dessous :



où :

- $X$ ,  $S$ ,  $A$ ,  $O$  et  $C$  sont respectivement la biomasse (bactéries *E. coli*), le substrat (ici le glucose), l'acétate, l'oxygène et le dioxyde de carbone,
- les paramètres  $k_j$  pour  $1 \leq j \leq 10$  sont les coefficients pseudo-stœchiométriques. Dans ce qui suit, ils sont nommés « coefficients stœchiométriques »,
- $r_i$  sont les vitesses de réaction et sont associées à la biomasse pour  $1 \leq i \leq 3$ .

Chaque réaction décrit une voie catabolique<sup>2</sup> spécifique. La première réaction correspond à l'oxydation du glucose. Elle est active sous des conditions de fonctionnement en mode aérobie, c'est-à-dire en présence d'oxygène. Cette réaction décrit l'oxydation d'une partie du glucose en dioxyde de carbone en utilisant l'oxygène comme accepteur final. C'est une réaction exergonique : elle dégage en se réalisant une forme d'énergie qui permet la biosynthèse des constituants cellulaires.



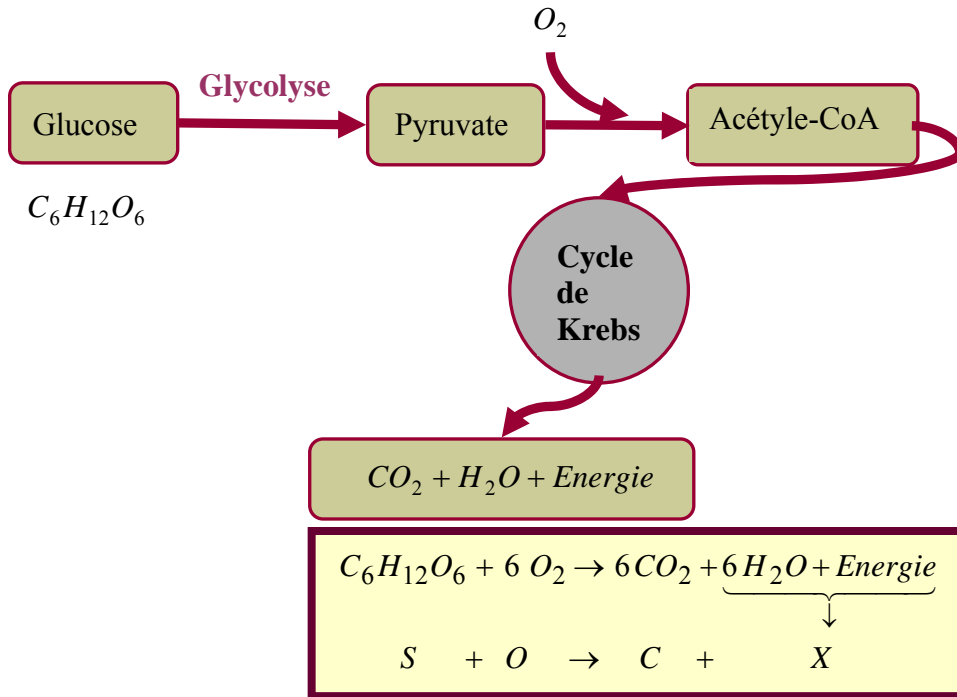
**Figure 2.3.** La glycolyse ou la dégradation du glucose.

Le glucose se dégrade en pyruvate par l'intermédiaire de la glycolyse, comme cela est décrit dans la Figure 2.3, puis la transformation du pyruvate en Acétyl-CoA permet l'activation du

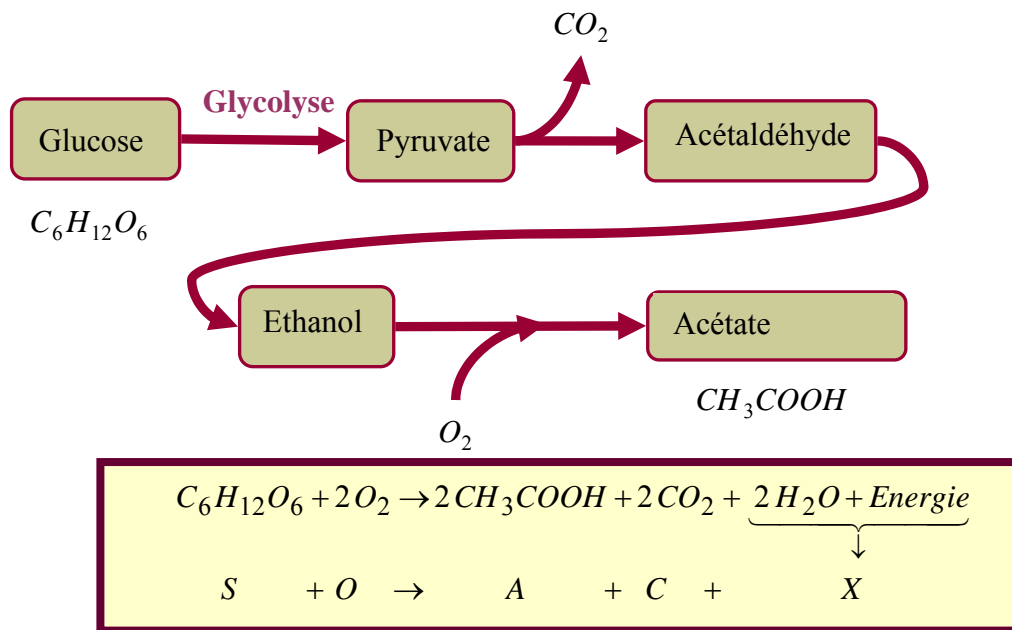
<sup>1</sup> Le métabolisme est l'ensemble des transformations moléculaires et des transferts d'énergie qui se déroulent de manière ininterrompue dans la cellule ou l'organisme vivant. C'est un processus ordonné, qui fait intervenir des processus de dégradation (catabolisme) et de synthèse organique (anabolisme).

<sup>2</sup> Le catabolisme est l'ensemble des réactions de dégradations moléculaire de l'organisme considéré. Il est le contraire de l'anabolisme, ensemble des réactions de synthèse. Les réactions de catabolisme sont des oxydations (ou des déshydrogénations) et sont exoénergétiques (produisant de l'énergie).

cycle de Krebs<sup>3</sup> (Figure 2.4). Ce dernier dégage du dioxyde de carbone, de l'eau et de l'énergie qui permet la reproduction des bactéries, et donc la croissance de la biomasse.



**Figure 2.4.** Métabolisme de l'oxydation complète du glucose.

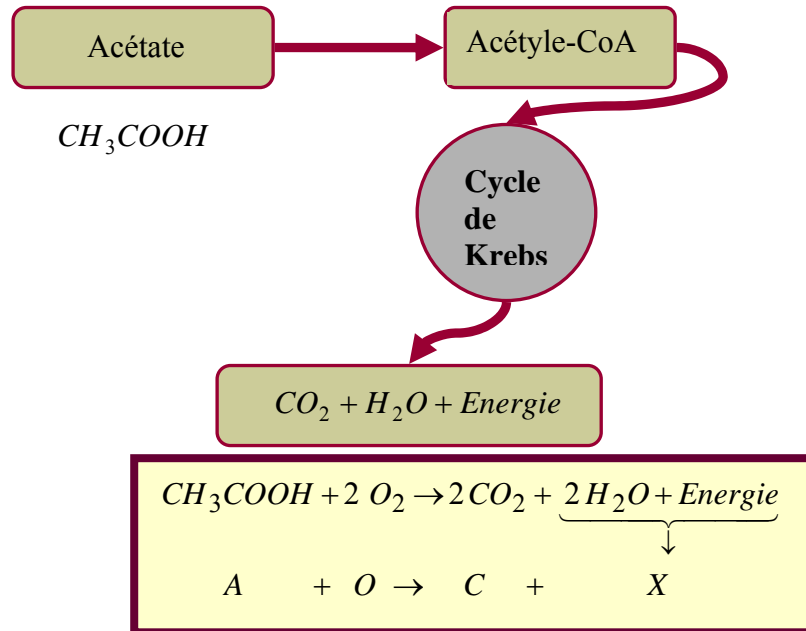


**Figure 2.5.** Métabolisme de fermentation du glucose.

<sup>3</sup> Voie métabolique cyclique, commune au catabolisme oxydatif final des glucides, des lipides et des acides aminés et fonctionnant dans les mitochondries des cellules aérobies. Au cours de ce cycle, l'acétyl-coenzyme A provenant de la dégradation des glucides, des lipides et des acides aminés s'unit à l'acide oxalo-acétique pour fournir de l'acide citrique. Au terme d'une série de réactions, l'acide oxalo-acétique est régénéré. Le bilan global est une libération de deux molécules de  $CO_2$  et de quatre paires d'atomes d'hydrogène et la libération d'énergie sous forme d'ATP (Adénosine triphosphate).

La deuxième réaction décrit la fermentation du glucose (Figure 2.5). Elle a lieu soit en conditions anaérobies, soit en présence d'excès de glucose (théorie du « goulot d'étranglement », Figure 2.7). Le glucose se transforme en éthanol par fermentation alcoolique et l'oxygène présent dans l'air permet l'oxydation de l'éthanol en acétate.

Enfin, la troisième réaction suit le même raisonnement que l'oxydation du glucose : l'acétate se dégrade en acétyl-CoA et ce dernier permet l'activation du cycle de Krebs (Figure 2.6).



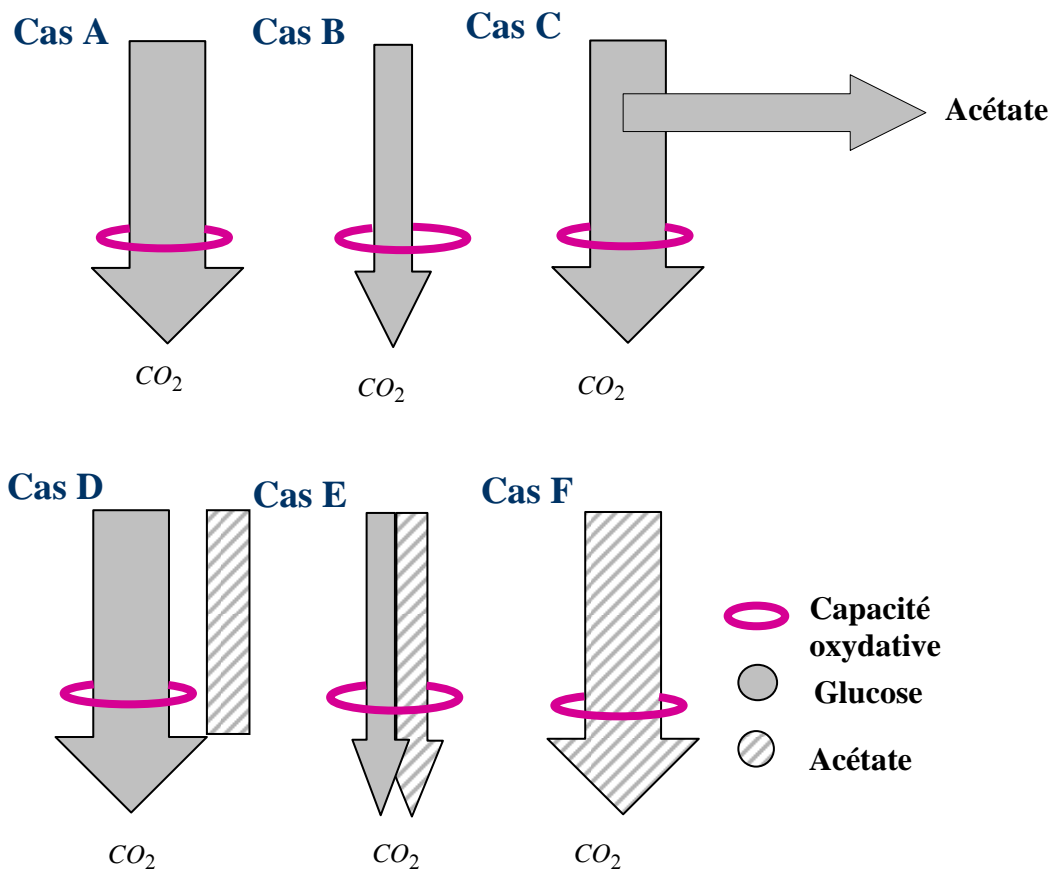
**Figure 2.6.** Métabolisme d'oxydation de l'acétate.

Ces trois réactions sont largement détaillées dans [Ko *et al.*, 1993] et [Ko *et al.*, 1994].

Pour compléter les points précédents, il est nécessaire de détailler dans notre cas la théorie du goulot d'étranglement. La Figure 2.7 présente une synthèse des travaux réalisés concernant la formation aérobie de l'acétate [Xu *et al.*, 1999] et [Luli et Strohl, 1990]. Elle décrit le concept de saturation développé dans ces travaux. Six cas de figure sont présentés :

- Cas A : la capacité oxydative <sup>4</sup>est complètement traversée par le glucose présent dans le milieu de culture,
- Cas B : la quantité de glucose est inférieure à la capacité oxydative,
- Cas C : la quantité de glucose est supérieure à la capacité oxydative. La quantité restante est transformée en acétate suivant une autre voie métabolique,
- Cas D : le glucose utilise toute la capacité oxydative malgré la présence de l'acétate dans le milieu de culture,
- Cas E : la somme des quantités de glucose et d'acétate est inférieure à la capacité oxydative,
- Cas F : le milieu de culture ne contient pas de glucose, l'acétate utilise donc toute la capacité oxydative.

<sup>4</sup> L'expression de la capacité oxydative est donnée plus loin.



**Figure 2.7.** Représentation du concept du « goulot d'étranglement » (Bottleneck theory [Galvanauskas *et al.*, 1998]) dans le cas de la culture des bactéries *E. coli*.

Ainsi, malgré la présence d'oxygène, une forte concentration en glucose inhibe la voie catabolique oxydative et favorise la production d'acétate. Lorsque le flux de glucose consommé par la cellule ne peut plus être entièrement oxydé (la capacité oxydative est saturée), le surplus est dirigé vers la voie catabolique fermentative. Ceci permet de définir deux régimes de fonctionnement dans le cas de culture des *E. coli* :

- Régime oxydatif défini par les deux voies cataboliques (2.10) et (2.12),
- Régime oxydo-fermentatif défini par les deux voies cataboliques (2.10) et (2.11).

Ces deux régimes de fonctionnement sont détaillés par la suite lors de l'élaboration du modèle macroscopique de la culture de bactéries *E. coli* en mode fed-batch, et devront être pris en compte lors de l'élaboration de stratégies de commande, de la phase de détermination du profil optimal d'alimentation à la synthèse de la loi de commande proprement dite.

## 2.4.2 Modèle macroscopique

On s'intéresse dans nos travaux uniquement au bioprocédé opéré en mode fed-batch. Les équations différentielles du modèle sont obtenues à partir du schéma réactionnel (2.10), (2.11) et (2.12) en appliquant la formulation du modèle dynamique général (2.3) :

$$\frac{d}{dt} \begin{bmatrix} X \\ S \\ A \\ O \\ C \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix}}_K \underbrace{\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}}_{\varphi} - D \underbrace{\begin{bmatrix} X \\ S \\ A \\ O \\ C \end{bmatrix}}_{\xi} + \underbrace{\begin{bmatrix} 0 \\ S_{in} D \\ 0 \\ OTR \\ 0 \end{bmatrix}}_F - \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ CTR \end{bmatrix}}_Q \quad (2.13)$$

$$\frac{dW}{dt} = DW$$

où :

- $X$ ,  $S$ ,  $A$ ,  $O$  et  $C$  sont respectivement les concentrations en biomasse, glucose, acétate, oxygène et dioxyde de carbone et sont représentées en g/kg,
- $W$  est le poids du milieu de culture,
- $K$  est la matrice des coefficients stœchiométriques,
- $r_1$ ,  $r_2$  et  $r_3$  sont les vitesses de réaction des trois voies cataboliques (2.10), (2.11) et (2.12),
- $S_{in}$  représente la concentration en glucose dans le milieu d'alimentation,
- $D$  est le taux de dilution, rapport du débit d'alimentation  $F_{in}$  au poids du milieu de culture  $D = F_{in}/W$ . L'unité de  $D$  est en 1/h,
- $OTR$  est le taux de transfert d'oxygène de la phase gazeuse vers le milieu de culture,
- $CTR$  est le taux de transfert du dioxyde de carbone du milieu de culture vers la phase gazeuse.

Les vitesses de réaction  $r_i$  pour  $1 \leq i \leq 3$  sont liées à la croissance de la biomasse, elles sont proportionnelles aux taux spécifiques  $\mu_i$  [Bernard, 2002] par la relation :

$$r_i = \mu_i X \quad (2.14)$$

A partir des équations (2.13) et (2.14), le modèle macroscopique du bioprocédé *E. coli* s'écrit sous la forme :

$$\begin{aligned} \frac{dX}{dt} &= (\mu_1 + \mu_2 + \mu_3)X - DX \\ \frac{dS}{dt} &= (-k_1\mu_1 - k_2\mu_2)X - D(S - S_{in}) \\ \frac{dA}{dt} &= (k_3\mu_2 - k_4\mu_3)X - DA \\ \frac{dO}{dt} &= (-k_5\mu_1 - k_6\mu_2 - k_7\mu_3)X - DO + OTR \\ \frac{dC}{dt} &= (k_8\mu_1 + k_9\mu_2 + k_{10}\mu_3)X - DC - CTR \\ \frac{dW}{dt} &= DW \end{aligned} \quad (2.15)$$

où  $\mu_1, \mu_2$  et  $\mu_3$  sont les taux spécifiques des trois voies cataboliques (2.10), (2.11) et (2.12) liés à la cinétique réactionnelle avec comme unité 1/h .

### 2.4.3 Cinétique des *E. coli*

La modélisation de la cinétique réactionnelle des *E. coli* est largement étudiée dans la littérature. Ainsi [Kleman et Strohl, 1994] ont démontré que la capacité oxydative  $q_{OS}$  des *E. coli* est inhibée par l'acétate. En se basant sur ce résultat, l'expression de la capacité oxydative critique est donnée plus tard dans [Xu *et al.*, 1999] :

$$q_{OS\ crit} = q_{o\ max} \frac{k_{io}}{k_{io} + A} \quad (2.16)$$

où :

- $q_{o\ max}$  est le taux spécifique maximum de consommation d'oxygène,
- $k_{io}$  est la constante d'inhibition de l'oxygène.

Le taux spécifique  $q_S$  associé à la consommation du substrat (glucose) peut avoir plusieurs expressions. Le Tableau 2.1 résume les expressions rencontrées dans la littérature en précisant leurs sources.

Source	Expression du taux spécifique associé à la consommation du substrat
[Monod, 1942]	$q_S = q_{s\ max} \frac{S}{k_s + S}$ <p>où <math>q_{s\ max}</math> est le taux spécifique maximum de consommation de glucose, et <math>k_s</math> est la constante de saturation du glucose.</p>
[Galvanauskas <i>et al.</i> , 1998], [Xu <i>et al.</i> , 1999]	$q_S = q_{s\ max} \frac{S}{k_s + S} \frac{k_{is}}{k_{is} + A}$ <p>où <math>k_{is}</math> est la constante d'inhibition de l'acétate liée au taux <math>q_S</math>.</p>
[Cockshott et Bogle, 1992]	$q_S = q_{s\ max} \frac{S}{k_s + S} \left( 1 - \frac{A}{k_{is}} \right)$
[Alba et Calvo, 2000]	$q_S = q_{s\ max} \frac{S}{S + k_s + \frac{k_s}{k_{is}} A^2}$

**Tableau 2.1.** Différentes expressions du taux spécifique  $q_S$  rencontrées dans la littérature.

Dans [Rocha, 2003], une étude comparative des différentes expressions a été menée en considérant les erreurs entre les sorties du bioprocédé et du modèle pour chaque cas et en

mettant en œuvre pour l'identification une méthode de moindres carrés non-linéaires. Le critère présente une valeur minimale quand le taux spécifique  $q_S$  suit la loi de Monod. En se basant sur ce résultat, l'expression du taux spécifique associé à la consommation du substrat retenue par la suite est :

$$q_S = q_{s \max} \frac{S}{k_s + S} \quad (2.17)$$

La même stratégie est décrite dans [Rocha, 2003] pour le choix du taux spécifique associé à la production de l'acétate. La comparaison est faite pour trois expressions possibles présentées dans le Tableau 2.2.

Source	Expression du taux spécifique associé à la production de l'acétate
[Xu <i>et al.</i> , 1999]	$q_{AC} = q_{ac \max} \left( \frac{A}{A + k_a} \right)$ <p>où <math>q_{ac \max}</math> est le taux spécifique maximal de consommation d'acétate et <math>k_a</math> est la constante de saturation d'acétate.</p>
[Galvanauskas <i>et al.</i> , 1998]	$q_{AC} = q_{ac \max}$
[Rocha, 2003]	$q_{AC} = q_{ac \max} \frac{A}{k_a + A} \frac{k_{ia}}{k_{ia} + A}$ <p>où <math>k_{ia}</math> est la constante d'inhibition d'acétate.</p>

**Tableau 2.2.** Différentes expressions du taux spécifique  $q_{AC}$  rencontrées dans la littérature.

L'expression retenue dans [Rocha, 2003] pour le taux spécifique associé à la production de l'acétate est également celle que nous considérerons par la suite :

$$q_{AC} = q_{ac \max} \frac{A}{k_a + A} \frac{k_{ia}}{k_{ia} + A} \quad (2.18)$$

Cette expression décrit une loi de Monod complétée d'un terme décrivant l'inhibition par l'acétate.

La capacité oxydative  $q_{OS}$  est liée au taux spécifique associé à la consommation du substrat  $q_S$  par la relation :

$$q_{OS} = k_{os} q_S \quad (2.19)$$

où  $k_{os}$  est le rendement d'oxygène par rapport à la consommation du glucose.

De même, le taux spécifique d'oxygène utilisé pour l'oxydation de l'acétate est lié au taux spécifique associé à la production de l'acétate comme suit :

$$q_{OA} = k_{oa} q_{AC} \quad (2.20)$$

où  $k_{oa}$  est le rendement de l'oxygène par rapport à l'oxydation de l'acétate.

L'expression du taux spécifique critique lié à la consommation du glucose est déduite des relations (2.16) et (2.19) et s'écrit :

$$q_{S\ crit} = \frac{q_{O\ max}}{k_{OS}} \frac{k_{io}}{k_{io} + A} \quad (2.21)$$

Les expressions des taux spécifiques  $\mu_1$ ,  $\mu_2$  et  $\mu_3$  sont obtenues à partir du concept du « goulot d'étranglement » (Figure 2.7) selon les différentes possibilités.

- Pour les cas A, B et C, le milieu de culture contient uniquement du glucose, c'est-à-dire que  $S > 0$  et  $A = 0$ . D'une part :

$$A = 0 \Rightarrow q_{AC} = 0 \text{ et } \mu_3 = 0 \quad (2.23)$$

D'autre part, deux cas peuvent être différenciés :

- a. Cas où le glucose est inférieur ou égal à la capacité oxydative (Figure 2.7, cas A et B) :

$$q_{OS} \leq q_{OS\ crit} \Rightarrow \begin{cases} \mu_1 = q_S / k_1 \\ \mu_2 = 0 \end{cases} \quad (2.24)$$

- b. Cas où le glucose est supérieur à la capacité oxydative (Figure 2.7, cas C) :

$$q_{OS} > q_{OS\ crit} \Rightarrow \begin{cases} \mu_1 = q_{S\ crit} / k_1 \\ \mu_2 = (q_S - q_{S\ crit}) / k_2 \end{cases} \quad (2.25)$$

- Pour les cas D et E, le glucose et l'acétate sont présents dans le milieu de culture, donc  $S > 0$  et  $A > 0$  :

Le glucose est inférieur à la capacité oxydative,  $q_{OS} \leq q_{OS\ crit}$  :

- a. Si la somme des taux spécifiques d'oxygène utilisés pour l'oxydation du glucose  $q_{OS}$  et de l'acétate  $q_{OA}$  est supérieure à la capacité oxydative (Figure 2.7, cas D) :

$$q_{OS} + q_{OA} > q_{OS\ crit} \Rightarrow \begin{cases} \mu_1 = q_S / k_1 \\ \mu_2 = 0 \\ \mu_3 = (q_{OS\ crit} - q_{OS}) / (k_4 k_{oa}) \end{cases} \quad (2.26)$$

- b. Si la somme des taux spécifiques d'oxygène utilisés pour l'oxydation du glucose  $q_{OS}$  et de l'acétate  $q_{OA}$  est inférieure ou égale à la capacité oxydative (Figure 2.7, cas E) :

$$q_{OS} + q_{OA} \leq q_{OS\ crit} \Rightarrow \begin{cases} \mu_1 = q_S / k_1 \\ \mu_2 = 0 \\ \mu_3 = q_{AC} / k_4 \end{cases} \quad (2.27)$$

<sup>5</sup> D'après [Galvanauskas *et al.*, 1998],  $\mu_i = q_S / k_i, i=1,2$ , de même pour  $\mu_3$ .

Le glucose est supérieur à la capacité oxydative,  $q_{OS} > q_{OS\,crit}$ , alors on retrouve le cas (2.25).

- Pour le cas F de la Figure 2.7, on considère que le milieu contient de l'acétate uniquement ( $A > 0$  et  $S = 0$ ) :

$$q_S = 0 \Rightarrow \begin{cases} \mu_1 = 0 \\ \mu_2 = 0 \\ \mu_3 = q_{AC}/k_4 \end{cases} \quad (2.28)$$

Le Tableau 2.3 est un récapitulatif des expressions des taux spécifiques détaillées précédemment. Ces expressions dépendent du régime et/ou sous-régime de fonctionnement des bactéries *E. coli*.

Régimes de fonctionnement		Expression de $\mu_1$	Expression de $\mu_2$	Expression de $\mu_3$
<b>Régime oxydatif</b> $q_{OS} \leq q_{OS\,crit}$	<b>Sous-régime 1</b> $q_{OS} + q_{OA} \leq q_{OS\,crit}$	$\frac{q_S}{k_1}$	0	$\frac{q_{AC}}{k_4}$
	<b>Sous-régime 2</b> $q_{OS} + q_{OA} > q_{OS\,crit}$	$\frac{q_S}{k_1}$	0	$\frac{q_{OS\,crit} - q_{OS}}{k_4 k_{oa}}$
<b>Régime oxydo-fermentatif</b>	$q_{OS} \geq q_{OS,crit}$	$\frac{q_{S\,crit}}{k_1}$	$\frac{(q_S - q_{S\,crit})}{k_2}$	0

**Tableau 2.3.** Expressions des taux spécifiques selon le régime de fonctionnement.

Ainsi, le modèle décrit dans (2.15) avec les expressions des taux spécifiques  $\mu_1$ ,  $\mu_2$  et  $\mu_3$  présentées dans le Tableau 2.3 constitue un modèle de simulation du bioprocédé *E. coli*. Ce modèle macroscopique s'avère relativement complexe, très non-linéaire en particulier à cause des cinétiques réactionnelles, de nature hybride à cause des différents régimes. Dans l'objectif de développer une loi de commande, ce modèle va être simplifié lors du paragraphe ci-dessous.

## 2.5 Mise en œuvre d'un modèle simplifié du bioprocédé de culture de bactéries *E. coli*

Dans un souci de simplification du modèle présenté par (2.15) pour toutes les problématiques liées à la synthèse d'une loi de commande – analyse de sensibilité, identification, recherche de profil optimal ou calcul de la commande proprement dit – une structure de modèle simplifiée doit être élaborée. Elle servira de support pour la suite et sera nommée « modèle de commande ». Ceci est rendu possible d'une part parce qu'en pratique l'oxygène dissous est toujours régulé de manière à maintenir une quantité suffisante pour ne pas laisser la bactérie mourir. D'autre part, on remarque avec les relations (2.15) que les concentrations en oxygène

dissous et en dioxyde de carbone n'interviennent pas dans les équations des évolutions des autres variables d'état, en l'occurrence, la biomasse, le substrat (glucose), l'acétate et le poids.

Le modèle simplifié proposé se limite alors à quatre équations différentielles et s'écrit sous la forme suivante :

$$\begin{aligned}
 \frac{dX}{dt} &= r_X X - \frac{F_{in}}{W} X \\
 \frac{dS}{dt} &= -q_S X + \frac{F_{in}}{W} (S_{in} - S) \\
 \frac{dA}{dt} &= r_A X - \frac{F_{in}}{W} A \\
 \frac{dW}{dt} &= F_{in}
 \end{aligned}
 \tag{2.29}$$

où :

- $W$  est le poids du milieu de culture exprimé en kg,
- $F_{in}=DW$  est le débit d'alimentation en substrat (glucose dans ce cas) en kg/h,
- les taux  $r_X$  et  $r_A$ <sup>6</sup> dépendent du régime de fonctionnement et sont décrits dans le Tableau 2.4.

Taux	Régime oxydatif		Régime oxydo-fermentatif
	Sous-régime 1	Sous-régime 2	
$r_X$	$\frac{q_S}{k_1} + \frac{q_{AC}}{k_4}$	$\frac{q_S}{k_1} + \frac{k_{os}}{k_4} \frac{q_{Scrit} - q_S}{k_{oa}}$	$\frac{q_S - q_{S,crit}}{k_2} + \frac{q_{S,crit}}{k_1}$
$r_A$	$-q_{AC}$	$k_{os} \frac{q_S - q_{Scrit}}{k_{oa}}$	$k_3 \frac{q_S - q_{S,crit}}{k_2}$

**Tableau 2.4.** Expressions des taux  $r_X$  et  $r_A$  selon le régime de fonctionnement.

## 2.6 Conclusions

Ce chapitre a permis la détermination du modèle de bioprocédé de culture de bactéries *Escherichia coli* qui sera utilisé dans les chapitres suivants, soit sous sa forme complète pour la mise en œuvre d'un modèle de simulation, soit sous sa forme simplifiée pour la synthèse de lois de commande.

Une première étape s'est attachée à définir la structure générale du modèle macroscopique telle que proposée par Bastin et Dochain, en faisant intervenir le concept de schéma réactionnel et en présentant les lois classiques rencontrées pour décrire les cinétiques réactionnelles. Dans un deuxième temps, ces concepts généraux ont été particularisés pour l'obtention du modèle macroscopique du bioprocédé de culture de bactéries *E. coli* reprenant au final celui développé par [Rocha, 2003]. Une dernière partie a défini une forme simplifiée

<sup>6</sup> On rappelle que, d'après (2.15) et avec (2.29),  $r_x = \mu_1 + \mu_2 + \mu_3$  et  $r_A = k_3\mu_2 - k_4\mu_3$

de ce modèle en ne considérant que quatre variables d'état, qui servira de modèle de commande du bioprocédé de culture de bactéries *E. coli* pour les études suivantes, en l'occurrence l'analyse de sensibilité, l'identification, la recherche d'un profil optimal et l'application de la commande prédictive non-linéaire.

Dans tous les chapitres à venir, les simulations seront réalisées en se basant sur le bioréacteur utilisé par [Rocha, 2003] et en reprenant les valeurs numériques définies ou identifiées pour ce support. Les Tableaux 2.5 et 2.6 regroupent toutes les valeurs numériques des coefficients stœchiométriques et des coefficients cinétiques nécessaires à la simulation.

Paramètre	$k_1$	$k_2$	$k_3$	$k_4$	$k_5$	$k_6$	$k_7$	$k_8$	$k_9$	$k_{10}$
Valeur	3,164	25,220	10,900	6,382	1,074	11,89	6,098	1,283	19,01	6,576
Unité	g/kg	g/kg	g/kg	g/kg	g/kg	g/kg	g/kg	g/kg	g/kg	g/kg

**Tableau 2.5.** Valeurs des coefficients stœchiométriques du bioréacteur de [Rocha, 2003].

Paramètre	$q_{s \max}$	$k_s$	$k_{os}$	$q_{o \max}$	$k_{io}$	$q_{ac \max}$	$k_a$	$k_{oa}$	$k_{ia}$
Valeur	1,832	0,1428	2,020	0,7218	6,952	0,0967	0,5236	1,996	5,85
Unité	g/(kg.h)	g/kg	-	g/(kg.h)	g/kg	g/(kg.h)	g/kg	-	g/kg

**Tableau 2.6.** Valeurs des paramètres des taux spécifiques du bioréacteur de [Rocha, 2003].

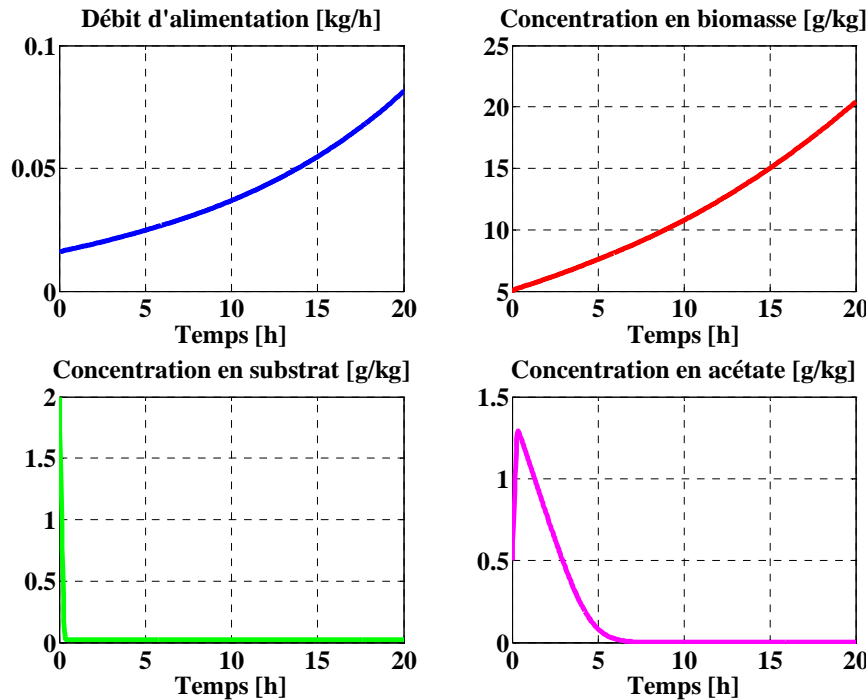
Pour conclure, en se basant sur cette configuration numérique et sur le modèle de simulation (2.15), les courbes ci-dessous ont pour objectif d'illustrer en boucle ouverte pour un scénario donné le fonctionnement du bioréacteur. Les conditions opératoires choisies sont les suivantes :

- conditions initiales :  $[X_0 \ S_0 \ A_0 \ O_0 \ C_0 \ W_0] = [5\text{g/kg} \ 2\text{g/kg} \ 0,5\text{g/kg} \ O_{sat} \ C_{sat} \ 3,17\text{kg}]$ ,
- profil d'alimentation  $F_{in} = 0,25 X W / S_{in}$ , avec ici  $S_{in} = 250\text{g/kg}$ . Ce profil est calculé dans [Rocha, 2003] pour obtenir une concentration en substrat constante. Ce type de profil est courant dans le cas des bioprocédés fonctionnant en mode fed-batch,
- temps de simulation est de 20 heures, ce qui correspond à la durée d'une culture en pratique.

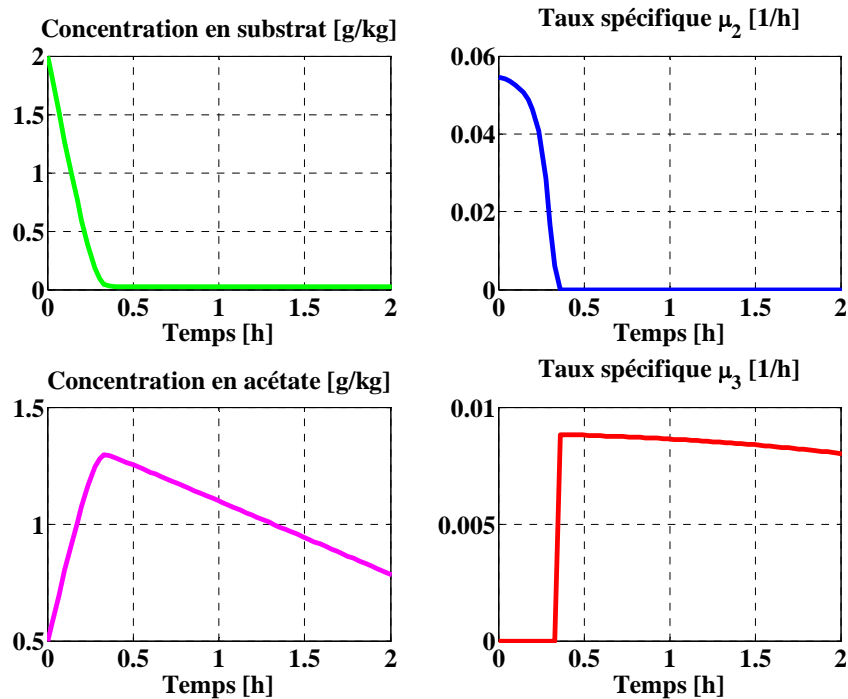
La Figure 2.8 montre les évolutions du débit d'alimentation, des concentrations en biomasse, en substrat et en acétate en boucle ouverte. Le profil d'alimentation ainsi que la concentration en biomasse ont une forme exponentielle, ce résultat est également courant dans le domaine des bioprocédés. Ce profil étant calculé pour une concentration en substrat constante, la figure montre qu'après quelques minutes, la concentration en substrat reste constante à une valeur très proche de zéro, une fois que ce dernier a été consommé.

La Figure 2.9 est un zoom sur les concentrations en substrat et en acétate et présente en plus, les évolutions des taux spécifiques  $\mu_2$  et  $\mu_3$ . En effet, comme le montre le Tableau 2.3, ces

taux spécifiques permettent de déterminer le régime de fonctionnement. Si  $\mu_2 = 0$ , le bioprocédé est en régime oxydatif et si  $\mu_3 = 0$ , il est en régime oxydo-fermentatif.



**Figure 2.8.** Evolution du débit d'alimentation, des concentrations en biomasse, en substrat et en acétate sur 20 heures.



**Figure 2.9.** Evolution des concentrations en substrat et en acétate et des taux spécifiques  $\mu_2$  et  $\mu_3$  sur 2 heures.

La figure montre que le bioprocédé démarre en régime oxydo-fermentatif, le glucose étant en quantité supérieure à la capacité oxydative ; on voit que la concentration en acétate augmente

et que la concentration en substrat diminue. Le glucose est donc consommé pour la production de la biomasse et de l'acétate, comme le décrivent les réactions (2.10) et (2.11) du schéma réactionnel. Une fois que le glucose est consommé, on remarque un changement de régime, le bioprocédé fonctionne en régime oxydatif, la concentration du substrat reste constante malgré le débit d'alimentation qui suit une exponentielle. Ceci signifie que le glucose est oxydé au fur et à mesure de l'ajout (réaction 2.10). La concentration en acétate décroît ce qui montre son oxydation selon la réaction (2.12).

## 3. Vers une stratégie d'identification

### 3.1 Introduction

Au chapitre précédent, un modèle simplifié, dit de commande, a été présenté pour les besoins de la synthèse de la loi de commande. Ce modèle est un modèle paramétrique qui nécessite donc l'identification de ses paramètres avant la mise en œuvre de la loi de commande.

Néanmoins, l'identification paramétrique s'avère difficile. En effet, les modèles des bioprocédés sont intrinsèquement non-linéaires et les paramètres du modèle peuvent ne pas être identifiables de façon indépendante. Quelques travaux ont eu lieu dans ce sens. [Chen et Bastin, 1996] introduisent la notion d'identifiabilité d'une matrice de paramètres. Ils présentent des théorèmes basés sur des relations algébriques déterminant l'identifiabilité des éléments de cette matrice. [Dochain et Vanrolleghem, 2001] présentent les notions d'identifiabilité structurelle et pratique. La première est basée sur la structure du modèle et la deuxième sur les expériences.

Malgré les différents problèmes d'identifiabilité rencontrés dans le cas des bioprocédés, quelques approches d'identification paramétrique existent dans la littérature. L'approche la plus courante consiste à déterminer les paramètres de manière non systématique en utilisant à la fois des données provenant de la littérature et le résultat de déterminations expérimentales partielles en mode batch, fed-batch et/ou continu [Alba et Calvo, 2000]. Une approche intéressante consiste à découpler la stœchiométrie de la cinétique en utilisant une transformation d'état afin d'identifier les coefficients stœchiométriques par la méthode de régression linéaire. Cette approche a été développée dans [Bastin et Dochain, 1990] et reprise deux ans plus tard dans [Chen, 1992]. Enfin, l'estimation simultanée de l'ensemble des paramètres peut être envisagée. Cette démarche est développée dans [Rocha, 2003] pour l'identification d'un modèle de culture des *E. coli*, où les régimes oxydatif et oxydo-fermentatif sont distingués.

Ce chapitre propose deux approches simples pour l'identification des paramètres du modèle de commande de la culture des *E. coli*. La première stratégie se base sur l'idée développée dans [Bastin et Dochain, 1990] qui consiste à découpler la stœchiométrie de la cinétique. Cette approche permet d'identifier les coefficients stœchiométriques à l'aide d'une régression linéaire, à condition que le modèle soit structurellement identifiable. En utilisant ces résultats, les paramètres cinétiques sont ensuite identifiés par une procédure de moindres carrés non-linéaires. La deuxième stratégie se base sur l'étude des sensibilités paramétriques du modèle (comme proposé dans [Rocha, 2003], [Pertev *et al.*, 1997] et [Smets *et al.*, 2002], pour le cas des *E. coli*, *Saccharomyces Cerevisiae* et *Azopirillum* respectivement). Cette étude permet de déterminer les paramètres les plus influents, selon le scénario de culture utilisé. L'identification paramétrique se fait alors en deux étapes et en utilisant deux scénarii différents. La première étape consiste à identifier les paramètres les plus influents pour le premier scénario. Ces résultats sont ensuite utilisés pour l'identification des paramètres restants et qui ont une influence plus significative pour le deuxième scénario.

Dans ce chapitre, les deux stratégies choisies sont développées, puis deux outils d'évaluation de la qualité des résultats d'identification paramétrique obtenus sont présentés. Elles sont ensuite appliquées à l'identification des paramètres du modèle de commande présenté dans le

chapitre 2 et comparées aux résultats d'une identification simultanée de tous les paramètres présentée dans [Rocha, 2003].

## 3.2 Identification basée sur le découplage

La stratégie d'identification paramétrique proposée ici repose sur l'idée d'identifier séparément les paramètres stœchiométriques et les paramètres de la cinétique (découpler la stœchiométrie et la cinétique). Cette procédure est basée sur une transformation d'état qui permet de formuler le modèle dynamique en deux sous-modèles distincts. Le premier sous-modèle dépend uniquement du schéma réactionnel et est indépendant de la structure cinétique. Il peut être linéairement paramétrisé et est utilisé pour l'identification des coefficients stœchiométriques au moyen de régressions linéaires. Ce découplage de la stœchiométrie et de la cinétique présente l'avantage de permettre l'identification de la stœchiométrie indépendamment de la cinétique. De plus, les coefficients stœchiométriques peuvent être identifiés même si les données expérimentales ne sont pas suffisamment informatives pour l'identification des paramètres cinétiques. Une fois les paramètres stœchiométriques identifiés, les paramètres cinétiques peuvent être déterminés par une approche de moindres carrés non-linéaires, en minimisant l'erreur entre les sorties mesurées et celles du modèle.

Dans cette partie, la méthode de découplage de la stœchiométrie et de la cinétique est développée. Afin d'identifier les paramètres stœchiométriques, la notion de « C-identifiabilité » est introduite. Enfin, la méthode utilisée pour l'identification des paramètres cinétiques est détaillée.

### 3.2.1 Identification des paramètres stœchiométriques

Afin d'identifier les paramètres stœchiométriques, la méthode de découplage proposée dans [Bastin et Dochain, 1990] est développée, puis les conditions d'identifiabilité structurelle de ces paramètres sont présentées.

#### 3.2.1.1 Méthode de découplage

La stœchiométrie peut être découplée de la cinétique si les hypothèses suivantes sont vérifiées :

- le nombre de réactions  $n_r$  est inférieur au nombre de constituants  $n$  :  $n_r \leq n$ ,
- tous les états sont mesurés,
- la structure de la matrice des coefficients stœchiométriques est connue a priori,
- chaque colonne de la matrice des coefficients stœchiométriques contient un 1 (ou -1) correspondant au constituant par rapport auquel la réaction a été normalisée.

Le modèle dynamique général défini dans [Bastin et Dochain, 1990] et présenté au chapitre 2, équation (2.3), peut s'écrire sous la forme suivante :

$$\frac{d\xi_a}{dt} = K_a r(\xi_a, t) - D\xi_a - Q_a + F_a \quad (3.1)$$

$$\frac{d\xi_b}{dt} = K_b r(\xi_b, t) - D\xi_b - Q_b + F_b \quad (3.2)$$

où :

- $(\xi_a, \xi_b)$ ,  $(Q_a, Q_b)$  et  $(F_a, F_b)$  sont des partitions des vecteurs  $\xi$ ,  $Q$  et  $F$  respectivement,
- $K_a$  et  $K_b$  représentent deux sous-matrices de la matrice des coefficients stœchiométriques  $K$  de rang  $n_K$ .

Il existe toujours une transformation d'état telle que :

$$z = B\xi_a + \xi_b \quad (3.3)$$

où  $B$  est une matrice de dimension  $(n - n_K) \times n_K$ . Elle est l'unique solution de l'équation matricielle :

$$BK_a + K_b = 0 \quad (3.4)$$

En considérant le changement de variable (3.3), le modèle (3.1)-(3.2) est équivalent à :

$$\frac{d\xi_a}{dt} = K_a r(\xi_a, t) - D\xi_a - Q_a + F_a \quad (3.5)$$

$$\frac{dz}{dt} = -Dz + B(F_a - Q_a) + (F_b - Q_b) \quad (3.6)$$

Il est à noter que la dynamique de la variable  $z$  est indépendante de la stœchiométrie et de la cinétique. A partir de cette représentation, on introduit les quantités auxiliaires à temps variant suivantes :

$$\alpha_a(t) = V(t)\xi_a(t) - V(0)\xi_a(0) + \int_0^t (F_{out}\xi_a + V(Q_a - F_a)) d\tau \quad (3.7)$$

$$\alpha_b(t) = V(t)\xi_b(t) - V(0)\xi_b(0) + \int_0^t (F_{out}\xi_b + V(Q_b - F_b)) d\tau \quad (3.8)$$

où  $V$  est le volume du milieu de culture et  $F_{out}$  est le débit d'alimentation de soutirage.

La relation suivante entre  $\alpha_a(t)$  et  $\alpha_b(t)$  peut être établie :

$$\alpha_b(t) = -B\alpha_a(t) \quad (3.9)$$

Souvent,  $K_a$  est une matrice carrée et régulière. Elle est de dimension  $n_K \times n_K$ . L'équation (3.4) devient alors :

$$B = -K_b K_a^{-1} \quad (3.10)$$

Dans ce cas, la relation (3.9) peut également s'écrire sous la forme :

$$\alpha_b(t) = (K_b K_a^{-1}) \alpha_a(t) \quad (3.11)$$

Soit  $\omega = [\omega_1, \dots, \omega_l]^T$  un vecteur regroupant les éléments non nuls de la matrice  $B$ . Il est de dimension  $l$  et ne dépend que des coefficients stœchiométriques. L'expression (3.11) peut s'écrire sous la forme d'une régression linéaire comme suit :

$$\alpha_b(t) = R^T(\alpha_a(t)) \omega \quad (3.12)$$

où  $R^T(\alpha_a(t))$  est une matrice de dimension  $(n - n_K) \times l$  et est fonction de  $\alpha_a(t)$ .

Ainsi, l'estimation de  $\omega$  est le résultat classique d'une régression linéaire et s'écrit :

$$\hat{\omega} = \left( \sum_{j=1}^{N_{mes}} R(\alpha_a(j)) R^T(\alpha_a(j)) \right)^{-1} \sum_{j=1}^{N_{mes}} R(\alpha_a(j)) \alpha_b(j) \quad (3.13)$$

où  $N_{mes}$  représente le nombre de mesures.

Puisque le vecteur  $\omega$  ne dépend que des coefficients stœchiométriques, son estimation permet l'identification de ces coefficients, à condition que la matrice  $B$  soit structurellement identifiable et que les coefficients stœchiométriques soient « C-identifiables ». Le paragraphe suivant présente les conditions que doit satisfaire la matrice  $B$  dans ce cas.

### 3.2.1.2 Conditions de la « C-identifiabilité »

La détermination de la « C-identifiabilité » est proposée dans [Chen et Bastin, 1996] à l'aide de relations algébriques simples à vérifier.

#### - Condition nécessaire et suffisante de « C-identifiabilité » :

Soit  $k^{(j)}$  le vecteur contenant les coefficients stœchiométriques à identifier au sein de la  $j^{\text{ème}}$  colonne de  $K$ .

$k^{(j)}$  est « C-identifiable » si et seulement si il existe au moins une partition  $K^T = \begin{bmatrix} K_a^T & K_b^T \end{bmatrix}$  où  $K_a$  est une matrice régulière et ne contient aucun élément de  $k^{(j)}$ .

#### - Corollaire :

Quand le nombre de réactions  $n_r \leq 2$ ,  $k^{(j)}$  est « C-identifiable » si et seulement si  $\dim(k^{(j)}) - 1 \leq n - n_r$ .

#### Remarque :

Quand le modèle n'est pas « C-identifiable », il est possible d'y remédier lorsque l'on dispose d'informations supplémentaires. Par exemple, on verra plus loin que dans le cas du bioprocédé *E. coli*, il faudra distinguer les deux régimes de fonctionnement pour vérifier la condition de « C-identifiabilité ».

## 3.2.2 Identification des paramètres cinétiques

Nous avons vu que la méthode de découplage proposée dans [Bastin et Dochain, 1990] et dans [Chen, 1992] nous permet d'identifier les coefficients stœchiométriques à l'aide des moindres carrés linéaires. Cette méthode nous permet d'éviter les difficultés rencontrées dans

le cas de problèmes d'optimisation non-linéaire (notamment la présence de minima locaux et sensibilité à l'initialisation).

Dans un deuxième temps, les paramètres cinétiques sont déterminés en fixant les coefficients stœchiométriques aux valeurs déterminées précédemment. Les paramètres cinétiques sont estimés en minimisant le critère quadratique suivant [Dochain et Vanrolleghem, 2001] :

$$J(\rho) = \sum_{i=1}^{N_{mes}} (y_i(\rho) - y_{mes_i})^T Q_i^{-1} (y_i(\rho) - y_{mes_i}) \quad (3.14)$$

où  $\rho$  est le vecteur des paramètres à identifier,  $y_{mes_i}$  et  $y_i(\rho)$  sont respectivement les vecteurs des  $N_{mes}$  valeurs mesurées et simulées via le modèle aux instants de mesure  $t_i$  ( $i = 1, \dots, N_{mes}$ ) et  $Q_i$  est la matrice de covariance des erreurs de mesure.

Ce problème d'optimisation est résolu par des méthodes de minimisation de critère quadratique non-linéaire (par exemple avec les algorithmes de Gauss-Newton ou de Levenberg-Marquardt [Fletcher, 1987]).

**Remarque :**

L'avantage de cette stratégie réside dans l'identification des coefficients stœchiométriques indépendamment de la cinétique, et dans le fait que l'estimation de ces coefficients est le résultat d'une régression linéaire. Malgré les avantages de cette méthode, elle est très sensible aux bruits de mesures. Des méthodes plus complexes existent dans la littérature qui traitent ce type de problèmes. [Bogaerts, 1999] utilise le principe du maximum de vraisemblance pour prendre en compte les bruits de mesures.

Le but ici est de présenter une solution d'identification alternative, basée sur l'analyse de sensibilité.

### 3.3 Identification basée sur l'analyse de sensibilité

La procédure d'identification développée ici est basée sur l'analyse de sensibilité des sorties du modèle vis-à-vis de ses paramètres. Le but de cette analyse est de déterminer les paramètres du modèle les plus influents pour chaque type d'expérience (batch, fed-batch...), et d'en déduire une stratégie d'identification pouvant s'avérer plus performante qu'une identification simultanée de l'ensemble des paramètres. Cette méthode s'apparente à une stratégie par planification d'expériences. Dans ce cas, les expériences sont déterminées par résolution d'un problème d'optimisation. En effet, les conditions opératoires et le profil d'alimentation sont déterminés pour maximiser la sensibilité du modèle vis-à-vis d'un paramètre à identifier.

Dans le cadre de cette thèse, le processus d'identification proposé ci-dessous est composé de plusieurs étapes, en considérant plusieurs scénarii expérimentaux (par exemple, tout d'abord une ou plusieurs expériences en mode batch, puis une ou plusieurs expériences en mode fed-batch). Pour chacun de ces scénarii, les paramètres les plus influents sont estimés et les valeurs des paramètres obtenues peuvent être exploitées à l'étape suivante. La démarche proposée ici peut être généralisée en utilisant plusieurs jeux de données (correspondant à des conditions expérimentales différentes) et plusieurs étapes successives (où les valeurs des

paramètres obtenues à une étape peuvent servir soit pour fixer la valeur de certains paramètres à l'étape suivante, soit pour initialiser l'identification d'autres paramètres. Finalement, on peut imaginer de libérer à nouveau l'ensemble des paramètres et procéder à une identification initialisée par les valeurs obtenues aux étapes précédentes.

La motivation première de la stratégie proposée, comparativement à une stratégie d'identification simultanée de l'ensemble des paramètres [Rocha, 2003], vient du fait qu'elle évite les problèmes de sur-paramétrisation ou de conditionnement lors de la résolution des problèmes d'optimisation pendant les phases d'identification. En effet, dans le cas où des paramètres possèdent des influences faibles pour un scénario donné, leurs valeurs identifiées ne sont pas précises et peuvent s'avérer éloignées des vraies valeurs. Ainsi, la démarche proposée a pour but de robustifier la procédure d'identification vis-à-vis des bruits de mesure et de l'influence de l'initialisation de l'algorithme.

### 3.3.1 Choix des paramètres à identifier

L'analyse de sensibilité permet de déterminer les paramètres les plus influents pour un scénario donné. Dans le cas des bioprocédés, les fonctions de sensibilité sont souvent linéairement dépendantes, ce qui donne lieu à des valeurs estimées de paramètres fortement corrélées. Pour éviter ce problème, nous proposons une étude des dépendances linéaires entre les paramètres à l'aide de la matrice de Gram. Ainsi, le choix des paramètres à identifier se fait en se basant sur ces deux études.

#### 3.3.1.1 Analyse de sensibilité

L'analyse des fonctions de sensibilité donne un profil temporel de la sensibilité du modèle. C'est une approche rigoureuse qui permet de mieux caractériser le modèle utilisé. Cette analyse permet également de guider le choix des protocoles d'identification, en concentrant les campagnes d'identification sur la détermination des paramètres les plus influents.

Les fonctions de sensibilité peuvent être définies comme étant la sensibilité du modèle pour des variations de paramètres, le long de la trajectoire nominale de fonctionnement [Munack, 1991]. Des exemples de calcul des fonctions de sensibilité peuvent être trouvés dans la littérature, par exemple dans [Perteve *et al.*, 1997] et [Smets *et al.*, 2002] appliqués respectivement à l'étude de la bactérie *Brasilense d'Azopirillum* et à celle de *Saccharomyces Cerevisiae*. Dans ces deux cas, après analyse de sensibilité, une simplification du modèle a été conduite en éliminant du modèle les paramètres les moins significatifs ou en fixant leurs valeurs numériques à des données issues de la littérature.

#### Définition : Fonctions de sensibilité

Soit  $x_i, i = \overline{1, n}$  les  $n$  variables d'état du modèle considéré et  $\theta_j, j = \overline{1, p}$  les  $p$  paramètres de ce modèle. Les fonctions de sensibilité du modèle vis-à-vis des paramètres sont définies par :

$\partial x_i / \partial \theta_j$ . Ces fonctions sont obtenues après intégration des équations  $\frac{d}{dt} \left( \frac{\partial x_i}{\partial \theta_j} \right)$ , en utilisant

l'égalité suivante :

$$\frac{d}{dt} \left( \frac{\partial x_i}{\partial \theta_j} \right) = \frac{\partial}{\partial \theta_j} \left( \frac{dx_i}{dt} \right) \quad (3.15)$$

Or, la dynamique des états est donnée par :

$$\frac{dx_i}{dt} = \mathfrak{F}_i(x, u, \theta) \quad (3.16)$$

où  $\mathfrak{F}_i$  représentent les  $n$  équations différentielles des variables d'état  $x_i$  et les  $\mathfrak{F}_i$  sont des fonctions non-linéaires. Ainsi, à partir des équations (3.15) et (3.16), il vient :

$$\frac{d}{dt} \left( \frac{\partial x_i}{\partial \theta_j} \right) = \frac{\partial \mathfrak{F}_i}{\partial \theta_j} + \sum_{k=1}^n \frac{\partial \mathfrak{F}_i}{\partial x_k} \frac{\partial x_k}{\partial \theta_j} \quad (3.17)$$

Ainsi, le calcul des fonctions de sensibilité revient à résoudre un système d'équations différentielles ordinaires de dimension  $n(p+1)$ .

Cependant, les paramètres du modèle sont susceptibles d'avoir des ordres de grandeur assez différents, engendrant des difficultés numériques lors de l'analyse de sensibilité du modèle vis-à-vis des paramètres. Pour remédier à ce problème, il est préférable de raisonner sur les fonctions de sensibilité normalisées données par  $\theta_j (\partial x_i / \partial \theta_j)$  [Huang et Yeh, 2007].

Pour chaque scénario d'expérience et/ou mode de fonctionnement, le calcul des fonctions de sensibilité normalisées permet de classer les paramètres selon leur influence et de déterminer les paramètres les plus influents. Seuls ces derniers sont identifiés selon la stratégie que nous proposons afin de faciliter la procédure d'identification paramétrique qui est très complexe dans le cas des bioprocédés.

### 3.3.1.2 Détermination des dépendances linéaires entre les paramètres

Pour réaliser une identification paramétrique fiable, il est important de vérifier s'il existe des dépendances linéaires entre les paramètres. Cette dépendance s'analyse à partir du déterminant de la matrice de Gram des fonctions de sensibilité normalisées.

La matrice de Gram des fonctions de sensibilité normalisées est définie comme suit [Luenberger, 1969] :

$$G = \begin{bmatrix} \frac{1}{T} \int_0^T S_{\theta_1}^2 dt & \frac{1}{T} \int_0^T S_{\theta_1} S_{\theta_2} dt & \dots & \frac{1}{T} \int_0^T S_{\theta_1} S_{\theta_{p-1}} dt & \frac{1}{T} \int_0^T S_{\theta_1} S_{\theta_p} dt \\ \frac{1}{T} \int_0^T S_{\theta_2} S_{\theta_1} dt & \frac{1}{T} \int_0^T S_{\theta_2}^2 dt & \frac{1}{T} \int_0^T S_{\theta_2} S_{\theta_3} dt & \dots & \\ \vdots & \vdots & \ddots & & \\ \frac{1}{T} \int_0^T S_{\theta_{p-1}} S_{\theta_1} dt & \vdots & & \frac{1}{T} \int_0^T S_{\theta_{p-1}}^2 dt & \\ \frac{1}{T} \int_0^T S_{\theta_p} S_{\theta_1} dt & \frac{1}{T} \int_0^T S_{\theta_p} S_{\theta_2} dt & \dots & \frac{1}{T} \int_0^T S_{\theta_p} S_{\theta_{p-1}} dt & \frac{1}{T} \int_0^T S_{\theta_p}^2 dt \end{bmatrix} \quad (3.18)$$

où :

- $S_{\theta_j}$  est la fonction de sensibilité normalisée pour chaque paramètre  $\theta_j$ ,
- $T$  est la durée de l'expérience.

Le déterminant de la matrice  $G$  est nul si les paramètres sont linéairement dépendants et est non nul sinon.

L'étude de sensibilité et celle des dépendances linéaires permet alors de déterminer les paramètres à identifier, à savoir les paramètres les plus influents et qui sont linéairement indépendants.

### 3.3.2 Procédure d'identification

La stratégie d'identification proposée est basée sur l'analyse de sensibilité du modèle par rapport à ses paramètres. L'idée consiste à hiérarchiser les paramètres selon leur influence sur le modèle et mettre au point une procédure d'identification utilisant cette information.

Ainsi, les paramètres les plus influents sont déterminés à l'aide d'une analyse de sensibilité. Puis, le déterminant de la matrice de Gram pour les paramètres retenus – c'est-à-dire les plus influents – est calculé afin de vérifier que ces paramètres sont linéairement indépendants. Ensuite, ces paramètres sont estimés en minimisant le critère (3.14). Les autres paramètres non identifiés sont fixés, soit à des valeurs issues de la littérature, soit aux valeurs estimées lors d'une étape d'identification précédente.

## 3.4 Validation de l'identification

Afin de quantifier la qualité des valeurs identifiées, une analyse a posteriori est effectuée, basée sur le calcul de la matrice d'information de Fischer et sur le calcul de l'erreur quadratique moyenne entre les sorties du modèle avec les valeurs identifiées et les sorties mesurées du système, sur un nouveau jeu de mesure.

### 3.4.1 Intervalles de confiance

Les fonctions de sensibilité jouent un rôle important dans l'évaluation de l'identifiabilité pratique des paramètres. Une propriété importante est que si les fonctions de sensibilité sont proportionnelles, la matrice de covariance des erreurs d'estimation est singulière : le modèle est dès lors non identifiable [Robinson, 1985].

Les intervalles de confiance sur les valeurs identifiées sont obtenus par le calcul de la matrice d'information de Fischer [Ljung, 1987], définie par :

$$F_{Fischer} = \sum_{i=1}^N \left( \frac{\partial y_i}{\partial \theta} \right)^T Q_i^{-1} \left( \frac{\partial y_i}{\partial \theta} \right) \quad (3.19)$$

où les termes  $\frac{\partial y_i}{\partial \theta}$  sont les fonctions de sensibilité des variables de sortie par rapport aux paramètres. L'intervalle de confiance à 95% est approché par  $\pm 2\sigma_i$  où les écarts-type  $\sigma_i$  sont les termes diagonaux de  $\sqrt{F_{Fischer}^{-1}}$ .

### 3.4.2 Erreur quadratique

Une validation classique des valeurs identifiées pour les paramètres consiste à les valider sur des jeux de données non utilisés lors de la phase d'identification. En effet, il est d'usage de diviser l'ensemble des mesures en deux sous-ensembles : un ensemble pour l'identification,

assez riche en points de mesure, et un autre sous-ensemble, pour valider les modèles identifiés.

Ainsi, à partir des modèles identifiés, la validation consiste à comparer les sorties du modèle obtenues avec les paramètres identifiés, notées  $y_{iden}$ , et les sorties mesurées, notées  $y_{mes}$ . On calcule alors l'erreur quadratique moyenne, notée  $E_q$ , qui représente l'espérance du carré des erreurs entre les sorties mesurées et les sorties du modèle et s'écrit comme suit (estimateur non biaisé) :

$$E_q = \sqrt{\frac{1}{(N_{mes} - 1)} \sum_{k=1}^{N_{mes}} (y_{iden}(k) - y_{mes}(k))^2} \quad (3.20)$$

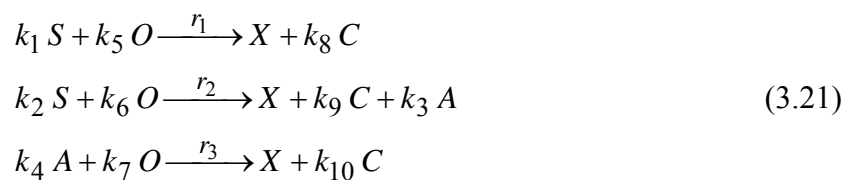
Cette erreur permet de quantifier la qualité des valeurs des paramètres identifiées. En effet, si cette erreur reste faible, elle valide le modèle identifié, et permet ainsi de détecter les cas de convergence vers des minima locaux (dus essentiellement aux bruits de mesure).

### 3.5 Application de la stratégie basée sur le découplage au bioprocédé *E. coli*

La stratégie développée à la section 3.2 et basée en partie sur la méthode de découplage entre stœchiométrie et cinétique est maintenant appliquée à l'identification paramétrique du bioprocédé *E. coli*. Dans un premier temps, l'étude de la possibilité de découpler la stœchiométrie de la cinétique est présentée, puis le découplage est réalisé pour chaque régime de fonctionnement. Enfin, les résultats de simulation sont présentés ainsi que la validation du modèle paramétrique ainsi obtenu.

#### 3.5.1 Découplage de la stœchiométrie et de la cinétique du bioprocédé

Avant de découpler la stœchiométrie de la cinétique, il est important de vérifier la « C-identifiabilité » du modèle. Reprenons le schéma réactionnel du bioprocédé *E. coli* présenté dans le chapitre 2 :



La matrice des coefficients stœchiométriques s'écrit donc sous la forme :

$$K = \begin{pmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{pmatrix} \quad (3.22)$$

Le rang de la matrice  $K$  est égal à 3. La condition de « C-identifiabilité » n'est pas vérifiée pour cette matrice car, pour tous les vecteurs colonnes de  $K$ , il est impossible de trouver une matrice  $K_a$  de dimension  $3 \times 3$  ne contenant pas d'éléments à identifier présents dans le vecteur colonne envisagé.

Pour remédier à ce problème, nous proposons de décomposer le modèle en deux sous-modèles, correspondants aux deux régimes de fonctionnement.

### ➤ Régime oxydatif

Le fonctionnement du bioprocédé en régime oxydatif correspond à l'oxydation du glucose et de l'acétate. Les réactions activées sont alors :



La matrice des coefficients stœchiométriques correspondante est notée  $K_O$  et s'écrit sous la forme :

$$K_O = \begin{pmatrix} 1 & 1 \\ -k_1 & 0 \\ 0 & -k_4 \\ -k_5 & -k_7 \\ k_8 & k_{10} \end{pmatrix} \quad (3.24)$$

Dans ce cas, le nombre de réactions est  $n_r = 2$ , on peut alors appliquer le corollaire présenté dans la section 3.2.1.2.

Rappelons que  $k^{(j)}$  représente le vecteur contenant les coefficients stœchiométriques au sein de la  $j^{\text{ème}}$  colonne de la matrice stœchiométrique. Les vecteurs  $k^{(j)}$  de la matrice  $K_O$  sont :  $k^{(1)} = [-k_1 \quad -k_5 \quad k_8]$  et  $k^{(2)} = [-k_4 \quad -k_7 \quad k_{10}]$ ,  $\dim(k^{(1)}) = \dim(k^{(2)}) = 3$ . En rappelant le nombre de constituants  $n = 5$ , la relation  $\dim(k^{(j)}) - 1 \leq n - n_r$  pour  $j = \{1, 2\}$  est vérifiée. Ainsi, la matrice  $K_O$  est « C-identifiable ». Dès lors, il est possible de découpler la stœchiométrie de la cinétique.

Nous proposons la partition suivante :

$$\xi_a = \begin{bmatrix} X \\ S \end{bmatrix} \quad \text{et} \quad \xi_b = \begin{bmatrix} A \\ O \\ C \end{bmatrix} \quad (3.25)$$

A partir de cette partition, les expressions des matrices  $K_a$  et  $K_b$  s'écrivent comme suit :

$$K_a = \begin{bmatrix} 1 & 1 \\ -k_1 & 0 \end{bmatrix} \quad \text{et} \quad K_b = \begin{bmatrix} 0 & -k_4 \\ -k_5 & -k_7 \\ k_8 & k_{10} \end{bmatrix} \quad (3.26)$$

La matrice  $K_a$  est inversible, donc il existe une unique matrice  $B$  telle que la relation (3.4) soit vérifiée. Ainsi, l'expression de la matrice  $B$  dans ce cas est :

$$B = -K_b K_a^{-1} = \begin{bmatrix} k_4 & \frac{k_4}{k_1} \\ k_7 & \frac{k_7 - k_5}{k_1} \\ -k_{10} & \frac{k_8 - k_{10}}{k_1} \end{bmatrix} \quad (3.27)$$

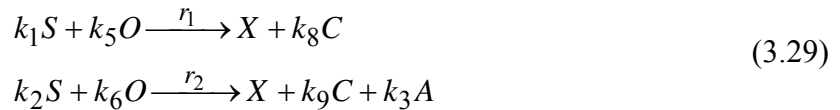
Dans ce cas, le vecteur  $\omega$  regroupant les éléments non nuls de la matrice  $B$  s'écrit sous la forme :

$$\omega = \left[ k_4 \quad k_7 \quad -k_{10} \quad \frac{k_4}{k_1} \quad \frac{k_7 - k_5}{k_1} \quad \frac{k_8 - k_{10}}{k_1} \right]^T \quad (3.28)$$

L'estimation du vecteur  $\omega$  par régression linéaire (3.13) permettra de déterminer les valeurs des coefficients stœchiométriques de la matrice  $K_O$ .

➤ **Régime oxydo-fermentatif**

Dans le cas du régime oxydo-fermentatif, les réactions du schéma réactionnel (3.21) activées sont :



La matrice des coefficients stœchiométriques correspondante est notée  $K_F$  et s'écrit sous la forme :

$$K_F = \begin{pmatrix} 1 & 1 \\ -k_1 & -k_2 \\ 0 & k_3 \\ -k_5 & -k_6 \\ k_8 & k_9 \end{pmatrix} \quad (3.30)$$

Le nombre de réactions ici est  $n_r = 2$ , le corollaire peut donc être utilisé comme pour le fonctionnement en régime oxydatif.

Les vecteurs  $k^{(j)}$  de la matrice  $K_F$  sont :  $k^{(1)} = [-k_1 \quad -k_5 \quad k_8]$  et  $k^{(2)} = [-k_2 \quad k_3 \quad -k_6 \quad k_9]$ ,  $\dim(k^{(1)}) = 3$  et  $\dim(k^{(2)}) = 4$ . La relation  $\dim(k^{(j)}) - 1 \leq n - n_r$  pour  $j = \{1, 2\}$  est ainsi vérifiée. La matrice  $K_F$  est donc « C-identifiable ».

Pour découpler la stœchiométrie de la cinétique, nous proposons la partition suivante :

$$\xi_a = \begin{bmatrix} X \\ A \end{bmatrix} \quad \text{et} \quad \xi_b = \begin{bmatrix} S \\ O \\ C \end{bmatrix} \quad (3.31)$$

Les expressions des matrices  $K_a$  et  $K_b$  correspondant à la partition (3.31) sont :

$$K_a = \begin{bmatrix} 1 & 1 \\ 0 & k_3 \end{bmatrix} \quad \text{et} \quad K_b = \begin{bmatrix} -k_1 & -k_2 \\ -k_5 & -k_6 \\ k_8 & k_9 \end{bmatrix} \quad (3.32)$$

Sachant que la matrice  $K_a$  est inversible, il existe une unique matrice  $B$  telle que :

$$B = -K_b K_a^{-1} = \begin{bmatrix} -k_1 & \frac{k_1 - k_2}{k_3} \\ -k_5 & \frac{k_5 - k_6}{k_3} \\ k_8 & \frac{k_9 - k_8}{k_3} \end{bmatrix} \quad (3.33)$$

Le vecteur  $\omega$  dans ce cas, s'écrit sous la forme suivante :

$$\omega = \begin{bmatrix} -k_1 & -k_5 & k_8 & \frac{k_1 - k_2}{k_3} & \frac{k_5 - k_6}{k_3} & \frac{k_9 - k_8}{k_3} \end{bmatrix}^T \quad (3.34)$$

On remarque que le vecteur  $\omega$  (3.34) contient 7 coefficients stœchiométriques mais qu'il est de dimension 6. Ainsi, l'estimation de ce vecteur ne permet pas d'identifier les 7 paramètres stœchiométriques. Les coefficients  $k_1$ ,  $k_5$  et  $k_8$  peuvent être identifiés directement à partir du vecteur  $\omega$  et les coefficients  $k_2$ ,  $k_6$  et  $k_9$  seront donnés en fonction de  $k_3$ .

### 3.5.2 Procédure d'identification paramétrique

L'identification des paramètres du bioprocédé *E. coli* dépend du régime de fonctionnement. L'étude est restreinte ici au fonctionnement en régime oxydo-fermentatif puisque seul ce régime sera considéré lors de la synthèse de la loi de commande comme il sera établi au chapitre 5. La démarche suivie dans ce qui suit est facilement transposable au cas oxydatif, en considérant les paramètres relatifs à ce régime.

On souhaite donc identifier les paramètres du modèle de commande présenté au chapitre 2, uniquement en régime oxydo-fermentatif.

Les équations différentielles des variables d'état en fonction des paramètres sont :

$$\begin{aligned} \frac{dX}{dt} &= \left( \frac{1}{k_1} \frac{q_{O\max}}{k_{Os}} \frac{k_{io}}{k_{io} + A} + \frac{1}{k_2} \left( q_{s\max} \frac{S}{k_s + S} - \frac{q_{O\max}}{k_{Os}} \frac{k_{io}}{k_{io} + A} \right) \right) X - \frac{F_{in}}{W} X \\ \frac{dS}{dt} &= -q_{s\max} \frac{S}{k_s + S} X + \frac{F_{in}}{W} (S_{in} - S) \\ \frac{dA}{dt} &= \frac{k_3}{k_2} \left( q_{s\max} \frac{S}{k_s + S} - \frac{q_{O\max}}{k_{Os}} \frac{k_{io}}{k_{io} + A} \right) X - \frac{F_{in}}{W} A \\ \frac{dW}{dt} &= F_{in} \end{aligned} \quad (3.35)$$

Ainsi, les paramètres à identifier ici sont :  $k_1$ ,  $k_2$ ,  $k_3$ ,  $q_{s\max}$ ,  $\frac{q_{o\max}}{k_{os}}$ ,  $k_s$  et  $k_{io}$ .

**Remarque :** La méthode de découplage est basée sur le schéma réactionnel, c'est la raison pour laquelle elle fait intervenir les coefficients stœchiométriques  $k_1$ ,  $k_2$ ,  $k_3$ ,  $k_5$ ,  $k_6$ ,  $k_8$  et  $k_9$ . Pour le modèle (3.35), nous avons besoin uniquement des valeurs de  $k_1$ ,  $k_2$  et  $k_3$ . Comme mentionné précédemment, l'avantage de cette méthode est qu'elle permet d'identifier ces paramètres à l'aide d'une régression linéaire en évitant tous les problèmes provenant de l'optimisation non-linéaire (minima locaux, sensibilité à l'initialisation...).

L'identification des paramètres du modèle (3.35) en utilisant la stratégie décrite dans la section 3.2 s'effectue en deux étapes :

1. Identification des paramètres stœchiométriques à partir de l'estimation du vecteur  $\omega$  (3.34). Comme montré au paragraphe précédent, dans ce cas, on obtiendra la valeur de  $k_1$  et une relation entre  $k_2$  et  $k_3$ .
2. Identification des paramètres cinétiques ( $q_{s\max}$ ,  $q_{o\max}/k_{os}$ ,  $k_s$ ,  $k_{io}$ ) et un paramètre stœchiométrique ( $k_2$  ou  $k_3$ ) en utilisant les résultats de la première étape. Cette identification est basée sur la résolution des moindres carrés non-linéaires.

Il est à noter que dans le cas du bioprocédé *E. coli*, le découplage stœchiométrie/cinétique devient en fait un découplage stœchiométrie/autres paramètres, puisqu'un paramètre stœchiométrique ( $k_2$  ou  $k_3$ ) est identifié lors de la deuxième étape (qui n'est donc pas uniquement une étape d'identification de paramètre cinétique).

**Remarque :** Il est évident que l'identification des paramètres  $q_{o\max}$  et  $k_{os}$  indépendamment n'est pas possible pour le modèle (3.35). Cette dépendance sera démontrée plus tard à l'aide de l'analyse de la matrice de Gram.

## 3.6 Application de la stratégie basée sur l'analyse de sensibilité au bioprocédé *E. coli*

La stratégie précédente utilise la méthode de régression linéaire qui est très sensible aux bruits de mesures. Une autre stratégie d'identification a été proposée à la section 3.3, reposant sur une approche de type moindres carrés non-linéaires. Cette stratégie basée sur l'analyse de sensibilité est appliquée ici afin d'identifier les paramètres du modèle du bioprocédé *E. coli*.

### 3.6.1 Détermination des paramètres à identifier

Pour déterminer les paramètres à identifier, nous allons suivre la procédure décrite dans la section 3.3.1. Cette stratégie est composée de deux étapes : premièrement, les fonctions de sensibilité sont calculées pour chaque mode de fonctionnement (batch et fed-batch) afin de déterminer les paramètres les plus influents. Deuxièmement, une analyse de la matrice de Gram est menée pour évaluer les dépendances entre les paramètres, dans le but de réduire la taille du vecteur de paramètres à identifier.

### 3.6.1.1 Fonctions de sensibilité du modèle du bioprocédé *E. coli*

Les expressions des fonctions de sensibilité du modèle de commande présenté au chapitre 2 sont déduites à partir des équations (3.15) et (3.17) comme suit :

$$\frac{d}{dt} \left( \frac{\partial x_i}{\partial \theta_j} \right) = \frac{\partial \mathcal{F}_i}{\partial \theta_j} + \frac{\partial \mathcal{F}_i}{\partial X} \frac{\partial X}{\partial \theta_j} + \frac{\partial \mathcal{F}_i}{\partial S} \frac{\partial S}{\partial \theta_j} + \frac{\partial \mathcal{F}_i}{\partial A} \frac{\partial A}{\partial \theta_j} + \frac{\partial \mathcal{F}_i}{\partial W} \frac{\partial W}{\partial \theta_j} \quad (3.36)$$

où :  $\theta_j = [k_1 \ k_2 \ k_3 \ k_4 \ q_{s \max} \ k_s \ k_{os} \ k_{io} \ q_{o \max} \ q_{ac \max} \ k_a \ k_{oa} \ k_{ia}]$  représente le vecteur des treize paramètres du modèle regroupant les coefficients stœchiométriques et cinétiques,  $j = \overline{1, 13}$ . L'indice des quatre variables d'état est :  $i = \overline{1, 4}$  et désigne les états ( $X, S, A, W$ ).

Puisque les expressions des  $\mathcal{F}_i$  dépendent du régime de fonctionnement du bioprocédé, les fonctions de sensibilité sont calculées en fonction du mode, régime et/ou sous-régime de fonctionnement.

Comme mentionné précédemment, on s'intéresse au calcul des fonctions de sensibilité normalisées pour pouvoir classer plus aisément les paramètres selon leur influence.

Compte tenu de la complexité du calcul analytique de l'équation (3.36), les fonctions de sensibilité normalisées,  $[\theta_j \cdot \partial \mathcal{F}_i / \partial \theta_j]$ , sont calculées numériquement par l'intermédiaire du programme CVodes [Serban et Hindmarsh, 2005] dans l'environnement *Matlab*<sup>TM</sup>.

Rappelons que le bioprocédé *E. coli* fonctionne selon deux régimes différents : le régime oxydo-fermentatif et le régime oxydatif qui se décompose quant à lui en deux sous-régimes. Pour chaque régime de fonctionnement, deux modes d'alimentation sont possibles, soit en mode batch, soit en mode fed-batch. Les expressions des fonctions  $\mathcal{F}_i$  dépendent du mode d'alimentation, des régimes et/ou des sous régimes de fonctionnement. L'analyse de sensibilité doit donc se faire pour tous ces cas (six en tout). Pour faciliter la lecture de ce mémoire, seule l'analyse de sensibilité en régime oxydo-fermentatif est présentée ici, les autres analyses sont détaillées en Annexe 1. En effet, la commande du bioprocédé sera réalisée principalement dans ce régime, comme il sera exposé au chapitre 5.

Il est à noter que lorsque le régime oxydo-fermentatif est actif, le coefficient stœchiométrique  $k_4$  et les coefficients cinétiques  $q_{ac \max}$ ,  $k_a$ ,  $k_{oa}$ ,  $k_{ia}$  n'interviennent pas dans le modèle. Les paramètres à identifier pour ce régime sont donc ceux du modèle (3.35), à savoir :  $k_1$ ,  $k_2$ ,  $k_3$ ,  $q_{s \max}$ ,  $q_{o \max}$ ,  $k_{os}$ ,  $k_s$  et  $k_{io}$ .

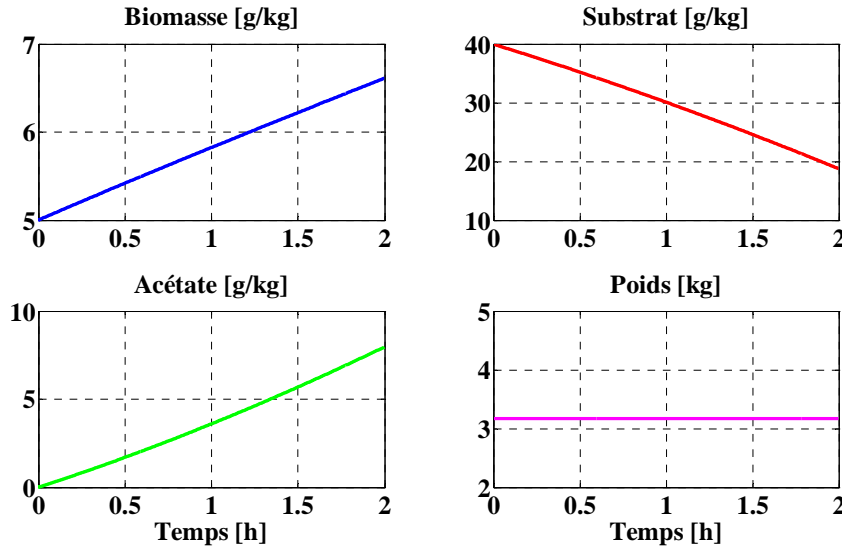
#### ➤ Bioprocédé fonctionnant en mode batch

Dans cette partie, le bioprocédé est supposé en mode batch et en régime oxydo-fermentatif. Dans ce cas, il n'y a pas d'alimentation en glucose et le poids du milieu de culture reste constant.

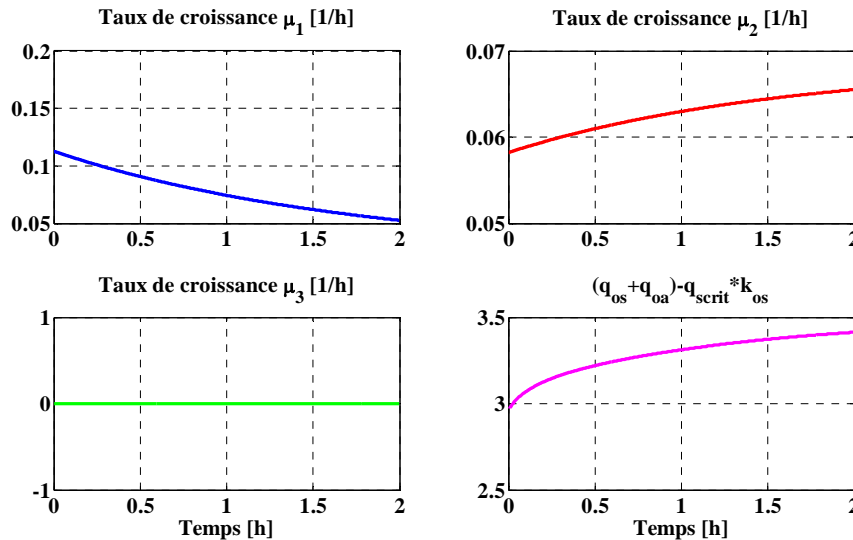
Pour analyser la sensibilité du modèle par rapport à ses paramètres, un scénario virtuel est choisi tel que, avec les paramètres donnés à la fin du Chapitre 2 :

- le vecteur initial  $[X_0 \ S_0 \ A_0 \ W_0] = [5\text{g/kg} \ 40\text{g/kg} \ 0\text{g/kg} \ 3,17\text{kg}]$
- l'intervalle de simulation considéré est  $[0 \ 2\text{h}]$  (horizon temporel choisi de façon à rester dans le régime oxydo-fermentatif).

Les Figures 3.1 et 3.2 représentent les évolutions des variables d'état du modèle et des taux de croissance et permettent de vérifier le type du mode et du régime choisis dans ce cas. En effet, le poids est constant sur l'intervalle de simulation, Figure 3.1, et le taux de croissance  $\mu_3$  est nul confirmant le régime de fonctionnement du bioprocédé, à savoir le régime oxydo-fermentatif. Notons par ailleurs que les conditions initiales ont été choisies pour placer le système dans le régime de fonctionnement souhaité.

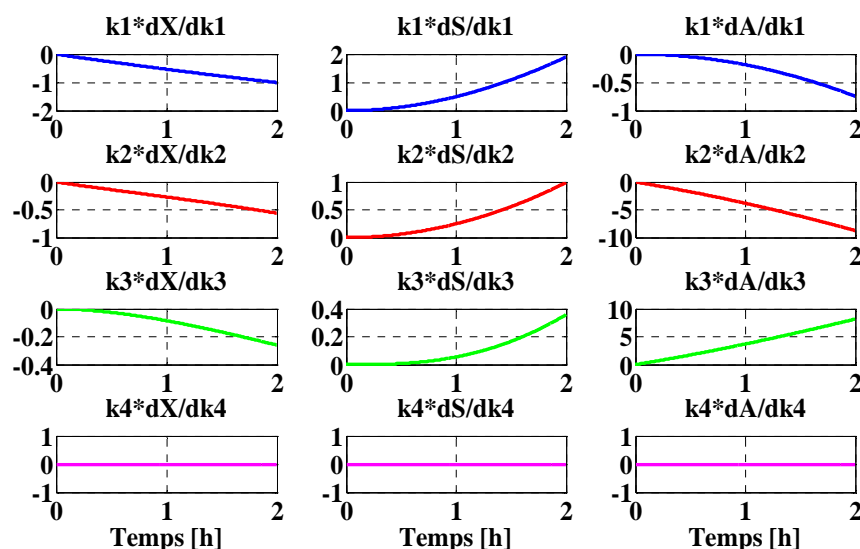


**Figure 3.1.** Evolution des variables d'état du modèle en mode batch et en régime oxydo-fermentatif sur l'intervalle [0 2h].



**Figure 3.2.** Evolution des taux de croissance en mode batch et en régime oxydo-fermentatif sur l'intervalle [0 2h].

Les Figures 3.3 et 3.4 représentent les fonctions de sensibilité normalisées des coefficients stœchiométriques et des coefficients cinétiques. Il est possible de vérifier que les paramètres  $k_4$ ,  $q_{ac\ max}$ ,  $k_a$ ,  $k_{oa}$  et  $k_{ia}$  n'ont pas d'influence sur le modèle, ce qui est prévisible comme mentionné précédemment. Par ailleurs, les huit paramètres restants possèdent des influences sur le modèle à différents degrés. Le Tableau 3.1 présente un classement de ces paramètres selon leur influence pour chaque variable d'état.



**Figure 3.3.** Fonctions de sensibilité normalisées des coefficients stœchiométriques pour chaque variable d'état en mode batch et en régime oxydo-fermentatif.

Influence sur la biomasse $X$	Influence sur le substrat $S$	Influence sur l'acétate $A$
$k_1$	$q_s \max$	$q_s \max$
$q_o \max$	$k_1$	$k_2$
$k_{os}$	$q_o \max$	$k_3$
$q_s \max$	$k_{os}$	$k_1$
$k_2$	$k_2$	$q_o \max$
$k_{io}$	$k_3$	$k_{os}$
$k_3$	$k_{io}$	$k_{io}$
$k_s$	$k_s$	$k_s$

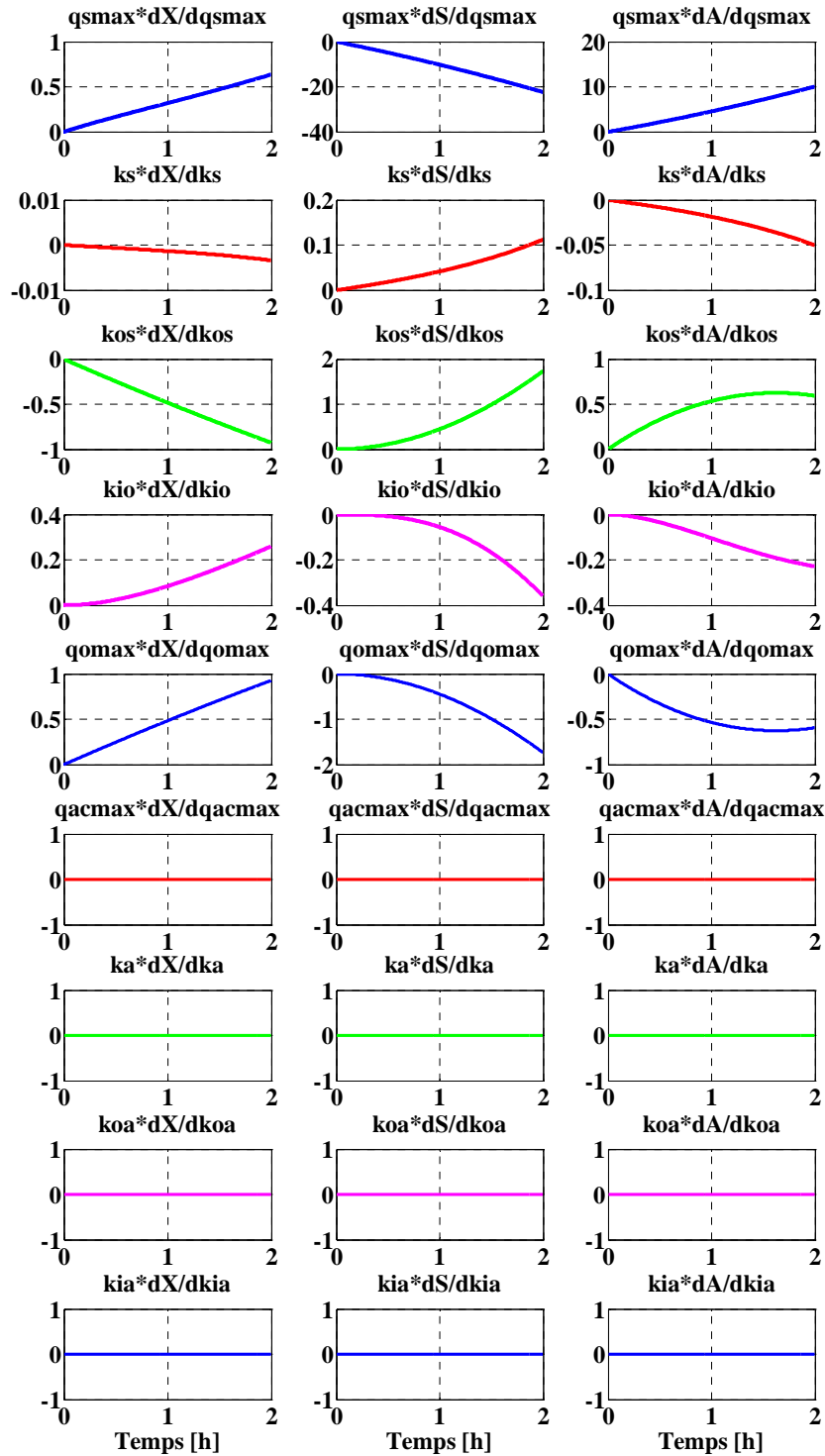
**Tableau 3.1.** Classement des paramètres selon leur influence pour chaque variable d'état (du plus influent au moins influent) en mode batch et en régime oxydo-fermentatif.

### ➤ Bioprocédé fonctionnant en mode fed-batch

Le scénario virtuel choisi dans ce paragraphe est celui d'une culture en régime oxydo-fermentatif, en mode fed-batch avec un débit d'alimentation en glucose constant et égal à 0,1 kg/h.

Les conditions opératoires choisies sont :

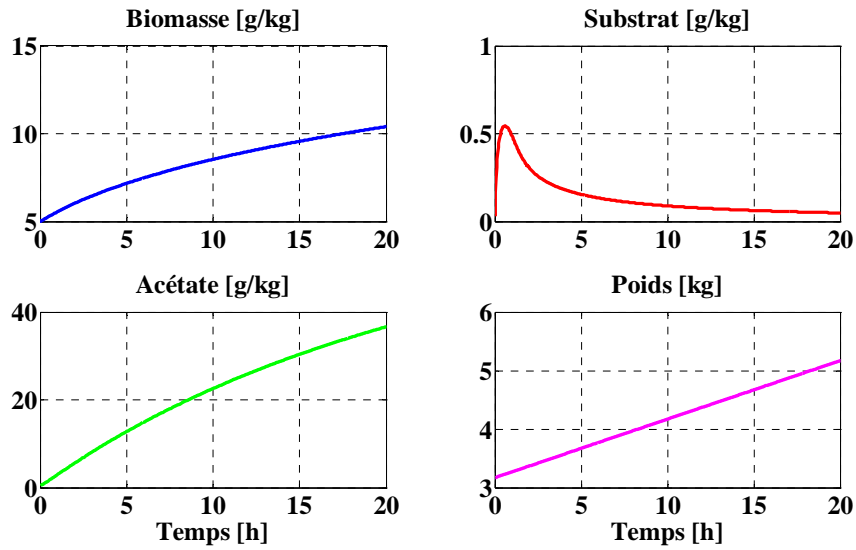
- le vecteur initial  $[X_0 \ S_0 \ A_0 \ W_0] = [5\text{g/kg} \ 0,0319\text{g/kg} \ 0,5\text{g/kg} \ 3,17\text{kg}]$ ,
- l'intervalle de simulation considéré est  $[0 \ 20\text{h}]$ .



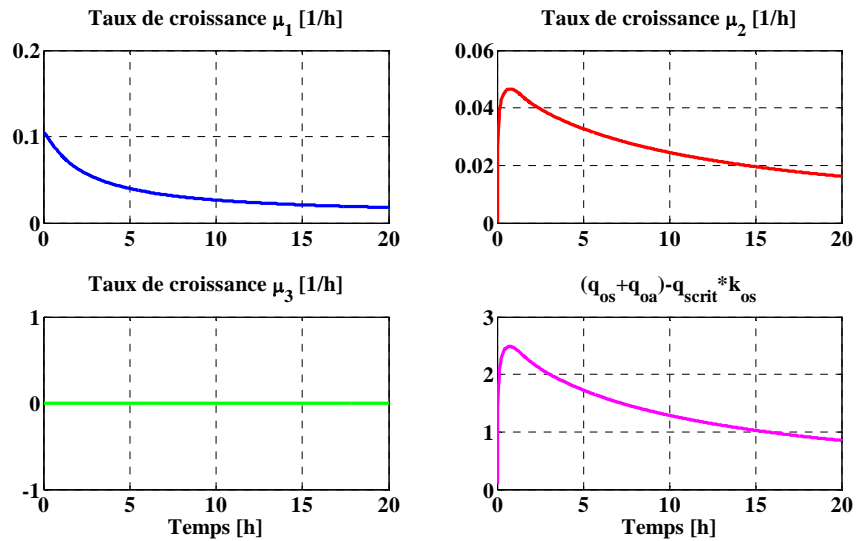
**Figure 3.4.** Fonctions de sensibilité normalisées des coefficients cinétiques pour chaque variable d'état en mode batch et en régime oxydo-fermentatif.

Les Figures 3.5 et 3.6 représentent les évolutions des concentrations en biomasse, en substrat et en acétate et le poids du milieu de culture, ainsi que les évolutions des taux de croissance. L'évolution linéaire du poids Figure 3.5 est due à l'alimentation en glucose supposée constante. Le taux de croissance  $\mu_3$  est nul sur l'intervalle [0 20h] ce qui prouve le fonctionnement en régime oxydo-fermentatif tout au long de cet intervalle. Comme précédemment, les états à l'instant initial ont été choisis tels que le système se place dans le

régime souhaité. D'autre part, du fait de l'alimentation en glucose, le système est maintenu dans le régime oxydo-fermentatif, d'où la possibilité de considérer un horizon temporel de simulation assez grand (de 20h, alors qu'il n'était que de 2h dans le cas précédent).



**Figure 3.5.** Evolution des variables d'état du modèle en mode fed-batch et en régime oxydo-fermentatif sur l'intervalle [0 20h].

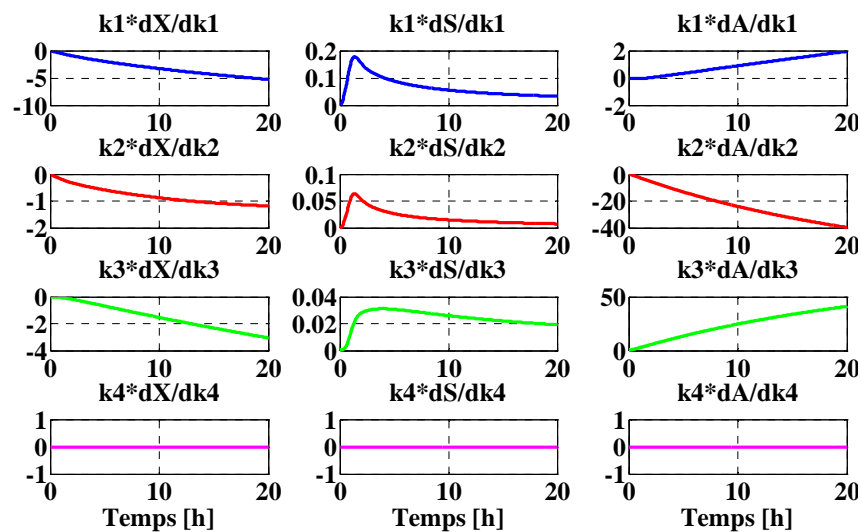


**Figure 3.6.** Evolution des taux de croissance en mode fed-batch et en régime oxydo-fermentatif sur l'intervalle [0 20h].

Les Figures 3.7 et 3.8 représentent les fonctions de sensibilité normalisées des coefficients stœchiométriques et des coefficients cinétiques. Comme pour le fonctionnement en batch, les paramètres  $k_4$ ,  $q_{ac\ max}$ ,  $k_a$ ,  $k_{oa}$  et  $k_{ia}$  n'interviennent pas, ce qui explique le comportement des fonctions de sensibilité normalisées Figure 3.7 pour  $k_4$  et Figure 3.8 pour  $q_{ac\ max}$ ,  $k_a$ ,  $k_{oa}$  et  $k_{ia}$ . L'influence des autres paramètres pour chaque variable d'état est résumée par le Tableau 3.2.

Influence sur la biomasse $X$	Influence sur le substrat $S$	Influence sur l'acétate $A$
$k_1$	$q_{s \max}$	$k_3$
$k_{os}$	$k_s$	$k_2$
$q_{o \max}$	$k_1$	$k_{os}$
$k_{io}$	$k_{os}$	$q_{o \max}$
$k_3$	$q_{o \max}$	$k_{io}$
$k_2$	$k_{io}$	$k_1$
$q_{s \max}$	$k_3$	$q_{s \max}$
$k_s$	$k_2$	$k_s$

**Tableau 3.2.** Classement des paramètres selon leur influence pour chaque variable d'état (du plus influent au moins influent) en mode fed-batch, en régime oxydo-fermentatif.



**Figure 3.7.** Fonctions de sensibilité normalisées des coefficients stœchiométriques pour chaque variable d'état en mode fed-batch et en régime oxydo-fermentatif.

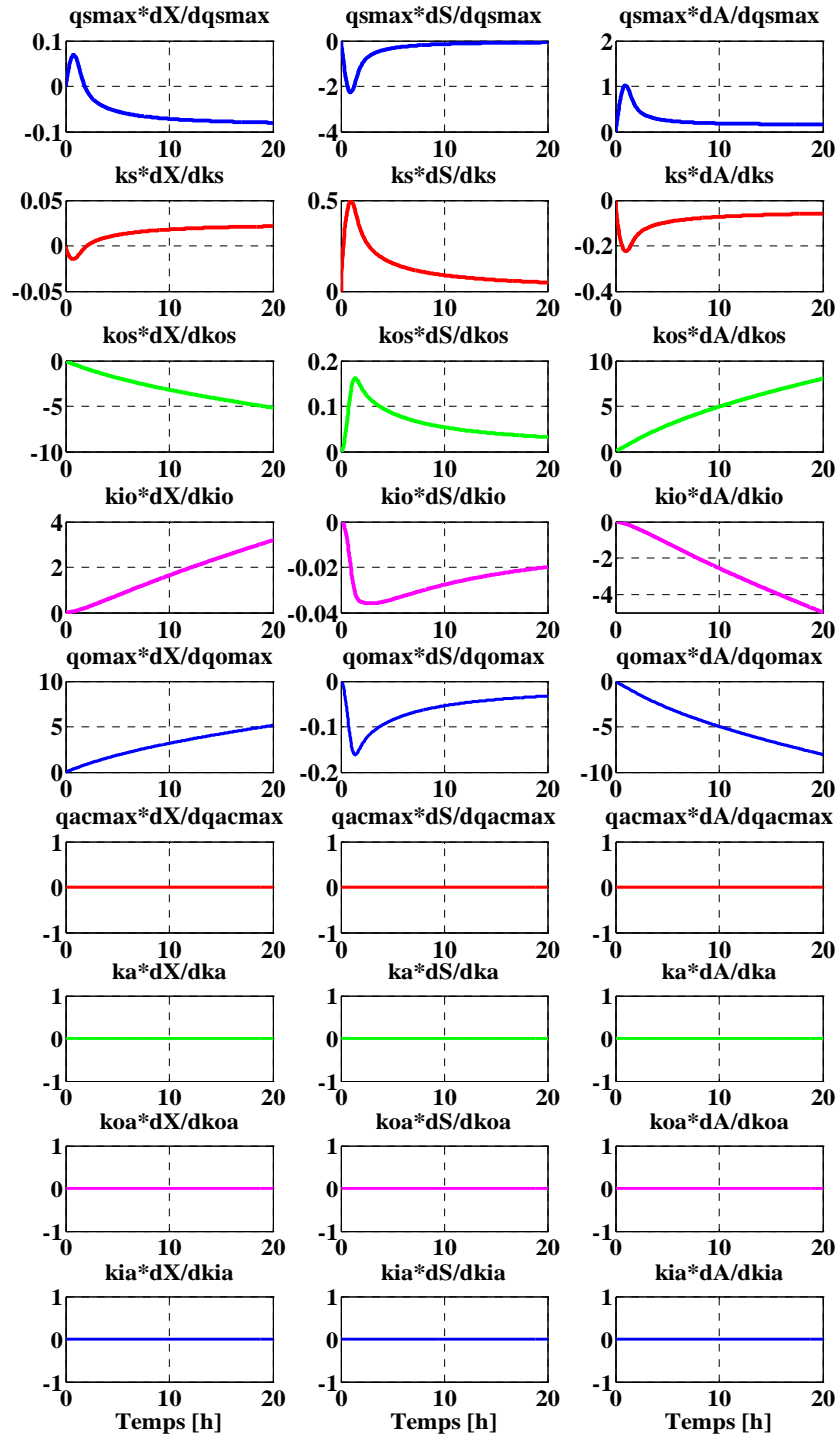
➤ **Récapitulatif**

L'examen des ordres de grandeur des fonctions de sensibilité permet alors de sélectionner, par mode de fonctionnement et par variable d'état les paramètres qui sont significativement les plus influents.

Ainsi en mode batch, à partir du Tableau 3.1, pour la sortie  $X$ , les paramètres  $k_1$ ,  $q_{o \max}$ ,  $k_{os}$  sont les plus significatifs. De façon similaire, on retient les paramètres  $q_{s \max}$ ,  $k_2$ ,  $k_3$  sur la sortie  $A$ . Pour la concentration en substrat  $S$ , le paramètre  $q_{s \max}$  est le plus influent et est le

seul à être sélectionné. En résumé, en tenant compte simultanément des trois variables d'état, le vecteur de paramètres à identifier se réduit pour le mode batch au vecteur  $[k_1 \ k_2 \ k_3 \ q_{s\max} \ q_{o\max} \ k_{os}]$ .

En appliquant la même démarche au mode fed-batch, en se basant sur le classement du Tableau 3.2, le vecteur de paramètres à identifier ne peut être réduit dans le cas où la totalité des variables d'état est considérée.



**Figure 3.8.** Fonctions de sensibilité normalisées des coefficients cinétiques pour chaque variable d'état en mode fed-batch et en régime oxydo-fermentatif.

### 3.6.1.2 Dépendances linéaires entre les paramètres

A partir des vecteurs de paramètres sélectionnés précédemment, il est nécessaire d'analyser la dépendance linéaire entre les divers paramètres retenus.

En mode batch, le calcul du déterminant de la matrice de Gram des fonctions de sensibilité normalisées associée aux paramètres  $q_{o\max}$  et  $k_{os}$  montre que ces deux paramètres sont linéairement dépendants. Cette dépendance est parfaitement visible par ailleurs dans les équations des taux de croissance (en régime oxydo-fermentatif, seul le rapport entre ces deux paramètres intervient dans la dynamique du système).

Dès lors, dans ce mode, le vecteur de paramètres à identifier se réduit pour le mode batch à 5 variables :  $(k_1, k_2, k_3, q_{s\max}, q_{o\max}/k_{os})$ . L'analyse des matrices de Gram pour les paramètres retenus pour chaque variable d'état permet de vérifier leurs indépendances. En effet, le calcul des déterminants des matrices de Gram des fonctions de sensibilité normalisées montre que :

- les paramètres  $k_1$  et  $q_{o\max}/k_{os}$  sont linéairement indépendants par rapport à la concentration en biomasse,
- les paramètres  $q_{s\max}$ ,  $k_1$  et  $q_{o\max}/k_{os}$  sont linéairement indépendants par rapport à la concentration en substrat,
- les paramètres  $q_{s\max}$ ,  $k_2$  et  $k_3$  sont linéairement indépendants par rapport à la concentration en acétate.

Par ailleurs, l'identification est réalisée en considérant les trois variables d'état simultanément. Ainsi, les paramètres  $(k_1, k_2, k_3, q_{s\max}, q_{o\max}/k_{os})$  considérés sont « globalement » linéairement indépendants (on dispose pour chacun d'eux de suffisamment d'information pour les identifier de façon indépendante).

En mode fed-batch, les déterminants de la matrice de Gram associée aux paramètres  $k_s$  et  $k_{io}$  sont non nuls pour les trois variables d'état. Ces paramètres sont donc linéairement indépendants et ce résultat sera utilisé dans la démarche d'identification.

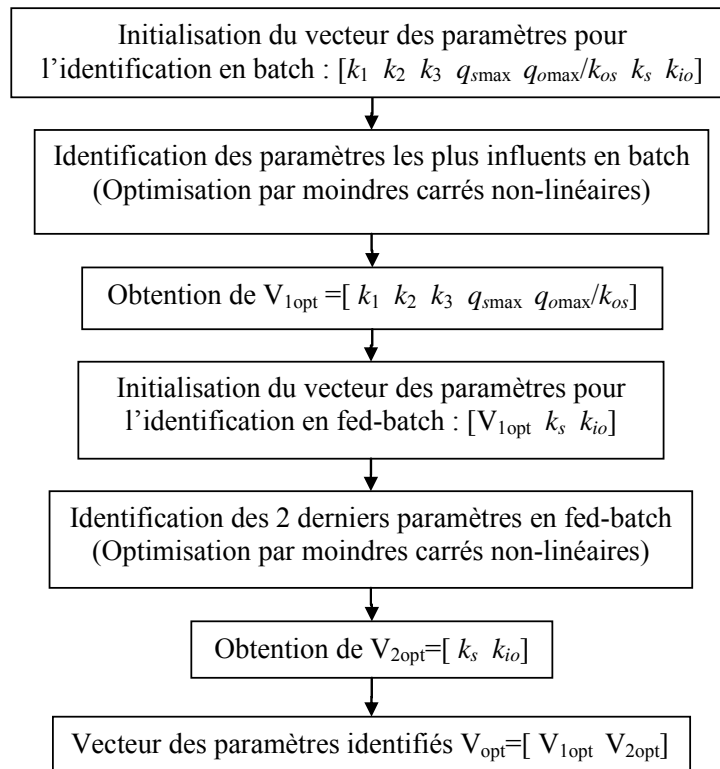
Plus globalement donc, le calcul de déterminants des matrices de Gram des fonctions de sensibilité normalisées permet de choisir les paramètres à identifier pour chaque étape et enrichit la connaissance des influences établies par les Tableaux 3.1 et 3.2.

### 3.6.2 Procédure d'identification paramétrique

Cette procédure utilise essentiellement les éléments mis en évidence lors des étapes précédentes. La première étape de la procédure d'identification proposée ici consiste à choisir les paramètres à identifier (dont on a montré l'indépendance), en se plaçant en mode batch, à savoir  $(k_1, k_2, k_3, q_{s\max}, q_{o\max}/k_{os})$ . Les valeurs des deux paramètres restants sont fixées à des valeurs issues de la littérature par exemple.

Une deuxième étape consiste à identifier les paramètres restants, à savoir  $k_s$ ,  $k_{io}$  (dont on a montré également l'indépendance) avec les mesures disponibles en mode fed-batch. Ce choix logique s'explique par leur influence significative dans ce mode de fonctionnement, illustré par les valeurs des fonctions de sensibilité normalisées. Les valeurs des paramètres

$k_1, k_2, k_3, q_{s\max}, q_{o\max}/k_{os}$  sont alors celles résultant de la première étape d'identification (en mode batch). La Figure 3.9 récapitule cette procédure d'identification.



**Figure 3.9.** Procédure d'identification basée sur l'analyse de sensibilité.

Cette procédure d'identification peut être répétée avec de nouveaux jeux de données, les valeurs identifiées à chaque étape pouvant servir soit pour fixer la valeur de certains paramètres à l'étape suivante, soit pour initialiser l'identification d'autres paramètres. Une autre possibilité serait d'utiliser les valeurs obtenues à la fin de cette procédure comme initialisation pour une résolution simultanée de l'ensemble des paramètres.

### 3.7 Résultats et validation en simulation

Les procédures d'identification décrites aux paragraphes 3.5.2 et 3.6.2 sont implantées dans l'environnement *Matlab*<sup>TM</sup>. Les valeurs des paramètres définies dans [Rocha, 2003] sont considérées ici comme des valeurs nominales. Elles permettent de valider la procédure d'identification en l'absence de données expérimentales réelles que nous n'avons pas pu obtenir à cause de différents problèmes techniques (voir la conclusion de ce mémoire). Pour cette même raison, les mesures sont générées à partir de simulations du modèle de commande (3.35) avec les valeurs nominales et en ajoutant du bruit de mesure sur chaque sortie (ici, le bruit de mesure est de 5% et la période d'échantillonnage des mesures est de 2 min). Le bruit choisi suit une loi normale. Dix expériences virtuelles sont générées, huit d'entre elles sont utilisées pour l'identification et les deux autres sont utilisées pour la validation de la qualité de l'identification (via le calcul de l'erreur quadratique). Les sorties considérées pour cette

étude sont : la concentration en biomasse  $X$ , la concentration en substrat  $S$  et la concentration en acétate  $A$ .

### 3.7.1 Stratégie basée sur le découplage

La procédure d'identification dans le cas de la stratégie basée sur le découplage a été détaillée au paragraphe 3.5.2.

L'estimation du vecteur (3.34) est obtenue à partir de l'expression (3.13) qui représente le résultat d'une régression linéaire.

Pour illustrer la première étape de cette stratégie, on supposera que les mesures ne sont pas bruitées (sinon, la stratégie s'avère mauvaise). Dans ce cas, le vecteur  $\omega$  obtenu est :

$$\omega = \begin{bmatrix} -k_1 & -k_5 & k_8 & \frac{k_1 - k_2}{k_3} & \frac{k_5 - k_6}{k_3} & \frac{k_9 - k_8}{k_3} \end{bmatrix}^T \quad (3.37)$$

$$= [-3,1640 \quad -1,0740 \quad 1,2830 \quad -2,0235 \quad -0,9923 \quad 1,6263]^T$$

Ainsi, la valeur identifiée du coefficient stœchiométrique  $k_1$  est égale à sa valeur nominale (ce qui est prévisible puisque le problème d'optimisation à résoudre est convexe). La relation entre les paramètres  $k_2$  et  $k_3$  est :

$$k_2 = -\omega(4)k_3 - \omega(1) \quad (3.38)$$

La deuxième étape de cette stratégie consiste à identifier les paramètres cinétiques ( $q_{s \max}$ ,  $q_{o \max}/k_{os}$ ,  $k_s$ ,  $k_{io}$ ) et le coefficient stœchiométrique  $k_3$ , la valeur de  $k_2$  sera ensuite déduite de la relation (3.38). Cette identification est le résultat de la minimisation du critère (3.14). L'optimisation de ce critère est réalisée avec un algorithme de Levenberg-Marquardt à l'aide du programme « *lsqnonlin* » de l'environnement *Matlab*<sup>TM</sup>. L'initialisation de cette minimisation est celle adoptée par [Rocha, 2003]. En pratique, l'initialisation se fera à partir de valeurs des paramètres issues de la littérature.

Les résultats de cette étape sont donnés par le Tableau 3.3. Les lignes notées 'borne inférieure' et 'borne supérieure' correspondent à l'intervalle de confiance à 95%.

Paramètres	$k_3$	$q_{s \max}$	$q_{o \max}/k_{os}$	$k_s$	$k_{io}$
Valeurs initiales	14,000	1,700	0,500	0,100	6,500
Valeurs nominales	10,900	1,832	0,357	0,142	7,142
Valeurs identifiées	<b>13,550</b>	<b>1,833</b>	<b>0,379</b>	<b>0,138</b>	<b>9,687</b>
Borne inférieure	0	1,800	0	0	0
Borne supérieure	84,670	1,886	1,080	0,715	54,850

**Tableau 3.3.** Identification des paramètres [ $k_3$   $q_{s \max}$   $q_{o \max}/k_{os}$   $k_s$   $k_{io}$ ] par la méthode par découplage.

Il en ressort que les valeurs identifiées sont proches des valeurs nominales. D'autre part, les valeurs nominales appartiennent à l'intervalle de confiance à 95%. Ceci constitue une première validation de ces résultats.

### 3.7.2 Stratégie basée sur l'analyse de sensibilité

La procédure d'identification pour cette stratégie a été détaillée au paragraphe 3.6.2. L'optimisation est également réalisée avec un algorithme de Levenberg-Marquardt à l'aide du programme « *lsqnonlin* » de l'environnement *Matlab*<sup>TM</sup>. Les résultats de cette stratégie sont résumés par les Tableaux 3.4 et 3.5. Comme nous l'avons précisé au paragraphe 3.7.1, les lignes notées 'borne inférieure' et 'borne supérieure' correspondent à l'intervalle de confiance à 95%. Les valeurs initiales sont celles considérées dans [Rocha, 2003].

Le Tableau 3.5 résume les résultats obtenus après la deuxième étape d'identification de la stratégie proposée. Les valeurs identifiées sont proches des valeurs nominales. On peut également vérifier que les valeurs nominales appartiennent bien à l'intervalle de confiance à 95%.

Paramètres	$k_1$	$k_2$	$k_3$	$q_{s \max}$	$q_{o \max}/k_{os}$
Valeurs initiales	2,000	20,000	14,000	1,700	0,500
Valeurs nominales	3,164	25,220	10,900	1,832	0,357
Valeurs identifiées	<b>2,834</b>	<b>25,460</b>	<b>10,830</b>	<b>1,828</b>	<b>0,324</b>
Borne inférieure	1,725	15,300	6,440	1,824	0,231
Borne supérieure	3,943	35,630	15,210	1,833	0,417

Tableau 3.4. Identification des paramètres [ $k_1$   $k_2$   $k_3$   $q_{s \max}$   $q_{o \max}/k_{os}$ ] en mode batch.

Paramètres	$k_s$	$k_{io}$
Valeurs initiales	0,100	6,500
Valeurs nominales	0,143	6,952
Valeurs identifiées	<b>0,142</b>	<b>7,142</b>
Borne inférieure	0,138	6,586
Borne supérieure	0,146	7,698

Tableau 3.5. Identification des paramètres [ $k_s$   $k_{io}$ ] en mode fed-batch.

### 3.7.3 Stratégie d'identification simultanée

Afin de comparer les résultats des deux stratégies précédentes à la stratégie d'identification simultanée de l'ensemble des paramètres et proposée dans [Rocha, 2003], le Tableau 3.6 présente les résultats dans ce cas. On remarque que les valeurs identifiées sont plus éloignées

des valeurs nominales et que l'écart-type est plus important. Malgré tout, une interprétation plus rigoureuse ne peut être effectuée qu'à partir de simulation sur de nouveaux jeux de données.

Paramètres	$k_1$	$k_2$	$k_3$	$q_s \max$	$q_o \max / kos$	$k_s$	$k_{io}$
Valeurs initiales	2,000	20,000	14,000	1,700	0,500	0,100	6,500
Valeurs nominales	3,164	25,220	10,900	1,832	0,357	0,143	6,952
Valeurs identifiées	<b>2,702</b>	<b>25,220</b>	<b>10,660</b>	<b>1,824</b>	<b>0,356</b>	<b>0,095</b>	<b>4,456</b>
Borne inférieure	1,747	0	0	1,785	0,181	0	0
Borne supérieure	3,656	66,200	29,990	1,862	0,531	0,765	16,690

Tableau 3.6. Identification globale des 7 paramètres en batch et fed-batch.

### 3.7.4 Validation des modèles identifiés

A partir des modèles identifiés dans ce qui précède, une validation sur de nouveaux jeux de données est effectuée. Pour ce faire, on considère des simulations du bioprocédé avec des bruits de mesure. Le Tableau 3.7 résume les différentes erreurs quadratiques pour les quatre modèles : modèle nominal où l'erreur est due uniquement au bruit de mesure, modèle issu d'une identification simultanée de l'ensemble des paramètres, modèle issu de l'application de la stratégie de découplage entre la stœchiométrie et la cinétique (section 3.5), et enfin modèle issu de la stratégie d'identification proposée dans ce travail de thèse (section 3.6).

Modèles	Nominal	Par approche Simultanée	Par approche par découplage	Par approche par analyse de sensibilité
Erreur quadratique sur X	0,0111	0,0378	0,0107	0,0181
Erreur quadratique sur S	0,0106	0,0275	0,0105	0,0155
Erreur quadratique sur A	0,0121	0,0184	0,0129	0,0167
Erreur quadratique totale	<b>0,0196</b>	<b>0,0503</b>	<b>0,0198</b>	<b>0,0291</b>

Tableau 3.7. Calcul des erreurs quadratiques.

La stratégie dite par découplage donne les meilleurs résultats (erreur quadratique de l'ordre du bruit de mesure). Or, cette stratégie n'a permis d'identifier qu'un seul paramètre ( $k_1$ ) dans la première phase avec une méthode de régression linéaire. Les autres paramètres ont été déterminés par une approche similaire à la stratégie par identification simultanée de l'ensemble des paramètres. La performance de cette stratégie s'explique par la non prise en compte du bruit de mesure dans l'identification du paramètre  $k_1$ . En effet, comme mentionné précédemment, il existe dans la littérature des méthodes appropriées pour gérer le bruit de mesure lors de la résolution d'un problème de régression linéaire, mais souvent complexes à

mettre en œuvre. Aussi, ici, le paramètre  $k_1$  identifié par la stratégie par découplage était égal à sa valeur nominale, d'où la bonne performance de cette méthode.

Ainsi, du Tableau 3.7, il ressort que la stratégie proposée consistant à classer les paramètres selon leur influence sur le modèle et à les identifier par étapes est celle qui donne le meilleur compromis entre qualité de la solution et complexité de l'algorithme. En effet, dans le cas de cette stratégie, aucune étude approfondie de la prise en compte du bruit de mesure n'a été développée, et seule l'analyse de sensibilité du modèle par rapport à ses paramètres est utilisée. Ainsi, cette stratégie est simple à appliquer, permettant d'obtenir des résultats corrects pour l'identification des paramètres du modèle.

### 3.8 Conclusions

Ce chapitre propose deux stratégies relativement simples pour l'identification paramétrique du modèle de commande du bioprocédé *E. coli*.

La première stratégie s'appuie sur la séparation de l'identification des paramètres stœchiométrique et des paramètres cinétiques (dite de découplage de la stœchiométrie et de la cinétique). Après vérification de la « C-identifiabilité » du modèle, la méthode de découplage permet d'identifier les paramètres stœchiométriques indépendamment de la cinétique à l'aide d'une régression linéaire. Puis les paramètres cinétiques et un paramètre stœchiométrique sont identifiés en utilisant la méthode des moindres carrés non-linéaires.

La première étape de cette stratégie étant très sensible aux bruits de mesures et pour éviter l'application de méthodes plus complexes, nous avons proposé une deuxième stratégie d'identification basée cette fois sur l'analyse de sensibilité. Cette étude ainsi qu'une analyse de la dépendance linéaire des paramètres permet de choisir les paramètres à identifier, selon le scénario d'expérimentation retenu. L'identification paramétrique se fait alors en deux étapes en fonction des résultats obtenus via l'analyse de sensibilité. Cette stratégie d'identification utilise également la méthode des moindres carrés non-linéaires.

La validation de la qualité des résultats obtenus en utilisant les deux stratégies est effectuée en évaluant la matrice d'information de Fischer et en calculant les erreurs quadratiques moyennes entre les sorties mesurées et les sorties du modèle identifié sur de nouveaux jeux de mesures. Des résultats de simulation sont présentés pour illustrer les deux stratégies proposées et leurs efficacités.

Pour conclure, il est à noter que la stratégie d'identification développée présente l'avantage d'être simple à implanter, et de donner une bonne approximation des valeurs des paramètres à identifier, de part la séparation des procédures d'identification des paramètres, pour éviter si possible les problèmes de conditionnement et de sur-paramétrisation. Elle est néanmoins sensible à l'initialisation des paramètres et à un niveau moindre risque de fournir des minima locaux du fait du bruit de mesures. La stratégie telle que décrite ci-dessus pourra être améliorée en y incorporant des procédures de robustification vis-à-vis de ces deux aspects.

En perspective, il reste encore à appliquer la démarche proposée à des données expérimentales, collectées à partir de cultures effectuées avec un bioprocédé à l'échelle de laboratoire de culture de bactéries *E. coli*.

## 4. Détermination d'un profil optimal d'alimentation

### 4.1 Introduction

La détermination d'un profil optimal d'alimentation d'un bioprocédé revient à résoudre un problème d'optimisation formulé à partir d'un critère exprimant le souhait de l'industriel selon le domaine d'activités. Globalement, deux types de problème existent : le critère est soit de maximiser la croissance de la biomasse, soit de maximiser la production d'un produit spécifique. En effet, dans le domaine pharmaceutique, la biomasse des bactéries *E. coli* est utilisée pour fabriquer des médicaments ou des vaccins selon la souche utilisée. La culture des levures *Saccharomyces Cerevisiae* peut être utilisée dans les domaines des biocarburants (production de l'Ethanol) et de l'agroalimentaire (biomasse utilisée pour fabriquer le pain, la bière, le vin...).

Le problème d'optimisation du fonctionnement d'un bioprocédé est abondamment traité dans la littérature scientifique (voir par exemple [Modak *et al.*, 1986], [Srinivasan *et al.*, 2003], [Smets *et al.*, 2004]). Lorsqu'un modèle de bioprocédé est parfaitement connu et suffisamment simple, le problème peut être résolu analytiquement en appliquant le principe du Maximum de Pontryagin [Bryson et Ho, 1969] ou l'approche simultanée [Tebhani *et al.*, 2006] ou encore d'autres méthodes connues pour la résolution des problèmes d'optimisation dynamique. Néanmoins, pour des modèles plus complexes, l'application de ces méthodes est difficile pour plusieurs raisons : présence de contraintes, temps de calcul très important, sensibilité à l'initialisation, présence de dynamiques différentes dans le système ...

Le but de ce chapitre est de déterminer un profil optimal d'alimentation du bioprocédé de culture de bactéries *E. coli* permettant de maximiser la quantité de biomasse  $X(t)W(t)$  à l'instant final  $t_f$ . Pour ce faire, un rappel des méthodes de résolution numérique des problèmes d'optimisation dynamique est présenté. Malheureusement, ces méthodes ne permettent pas de déterminer la solution optimale pour le problème considéré. La difficulté principale réside dans la nature du modèle du système. En effet, le système fonctionnant dans deux régimes de fonctionnement différents, il a de ce fait deux dynamiques différentes et possède donc une structure « hybride ». En conséquence, les méthodes d'optimisation classiques échouent du fait de la commutation d'un régime à un autre (ce qui augmente la sensibilité des algorithmes d'optimisation). Dans le cas où le modèle est simplifié, la solution obtenue est purement numérique [Rocha, 2004], et reste spécifique au cas d'étude traité (nécessité de recalculer la solution optimale selon les conditions opératoires).

[Srinivasan *et al.*, 2003] proposent une solution qui consiste à utiliser le principe du Maximum de Pontryagin non pas pour résoudre le problème d'optimisation mais pour caractériser la solution optimale. Cette démarche permet d'appliquer les méthodes directes plus facilement. Cette stratégie est mise en œuvre dans le cas de notre bioprocédé afin de déterminer les expressions analytiques des profils temporels optimaux qui maximisent une fonction coût à l'instant final.

Enfin, des résultats numériques permettant de déduire le profil optimal correspondant à la maximisation de la quantité de biomasse sont présentés.

## 4.2 Formulation du problème

Le problème d'optimisation à résoudre consiste à minimiser une fonction coût qui est une combinaison d'un coût terminal et d'un coût intégral. La formulation mathématique du problème d'optimisation est de la forme :

$$\min_{u(t)} J = \phi(x(t_f)) + \int_0^{t_f} L(x, u) dt \quad (4.1)$$

$$\text{sous les contraintes :} \quad \dot{x} = \mathcal{F}(x, u), \quad x(0) = x_0 \quad (4.2)$$

$$\psi(x, u) \leq 0, \quad \varphi(x(t_f)) \leq 0 \quad (4.3)$$

où :

- $J$  est le critère (fonction objectif) à minimiser,
- $x$  est le vecteur des variables d'état, de dimension  $n$ ,
- $x_0$  est le vecteur initial des variables d'état et est connu,
- $u$  est le vecteur des commandes, de dimension  $m$ ,
- $\psi$  est le vecteur des contraintes sur la trajectoire (incluant les contraintes sur les variables d'état et sur les commandes), de dimension  $n_\psi$ ,
- $\varphi$  est le vecteur des contraintes terminales, de dimension  $n_\varphi$ ,
- $\mathcal{F}$  est une fonction décrivant la dynamique du système,
- $\phi, L$  sont des fonctions scalaires représentant respectivement le coût terminal et le coût intégral,
- $t_f$  est le temps final, supposé fixé.

La résolution du problème d'optimisation (4.1) sous les contraintes (4.2) et (4.3) est largement répandue dans la littérature. De nombreux articles de synthèse témoignent de l'intérêt des méthodes pour la résolution de ce type de problème [Fletcher, 1991], [Biegler *et al.*, 2002], [Banga *et al.*, 2005], [Bonvin *et al.*, 2001] et [Srinivasan *et al.*, 2003]. On se basera par la suite sur l'état de l'art des approches de résolution proposé dans [Srinivasan *et al.*, 2003].

Principalement, les méthodes de résolution peuvent être classées en trois catégories :

- la résolution directe,
- la résolution indirecte (par application du Principe du Maximum),
- la résolution des équations de Hamilton-Jacobi-Bellman (HJB).

Dans ce chapitre, les deux premières approches seront présentées. La troisième méthode (résolution des équations HJB) est peu utilisée dans le domaine de la commande optimale des bioprocédés (difficulté de résolution) et de ce fait ne sera pas exposée. Son principe est détaillé dans [Bonnans *et al.*, 2001].

## 4.3 Résolution par l'approche directe

Cette approche consiste à transformer le problème d'optimisation (4.1)-(4.3) en un problème de programmation non-linéaire (dit PNL). Pour effectuer cette discrétisation du problème d'optimisation, deux approches différentes sont décrites dans la littérature :

- l'approche séquentielle,
- l'approche simultanée.

Dans ce qui suit, le principe de ces deux approches est brièvement rappelé.

### 4.3.1 Approche séquentielle

L'approche séquentielle est basée sur la paramétrisation du vecteur des commandes. Ainsi, la variable d'optimisation continue  $u$  est paramétrée en fonction d'un nombre fini de paramètres. L'algorithme de cette approche est comme suit :

1. Paramétrer les commandes en un nombre fini de variables de décision,
2. Choisir une initialisation des variables de décision,
3. Intégrer le système (4.2) jusqu'à l'instant final et calculer le critère  $J$  et les contraintes  $\psi$  et  $\varphi$ ,
4. Utiliser un algorithme d'optimisation pour mettre à jour les valeurs des variables de décision. Répéter les étapes 3 et 4 jusqu'à minimisation de la fonction objectif.

Le vecteur des commandes est généralement approché par des fonctions polynomiales. Les variables de décision sont donc les coefficients de ces polynômes. L'approximation la plus généralement utilisée est une commande constante par morceaux. Cette opération est appelée « Control Vector Parametrization (CVP) » dans la littérature, sa théorie est développée dans [Ray, 1981] et [Vassiliadis, 1993]. Elle est utilisée pour l'optimisation du fonctionnement des colonnes à distiller dans [Fikar *et al.*, 1998] et [Sorensen *et al.*, 1996].

Cependant, cette approche présente l'inconvénient d'être lourde en temps de calcul (notamment l'étape 3). En effet, le système est intégré exactement, même si les variables de décision sont loin de la solution optimale recherchée. Pour éviter cette intégration explicite, une autre approche est proposée, dite approche simultanée, et présentée ci-dessous.

### 4.3.2 Approche simultanée

Pour alléger les calculs, l'approche simultanée consiste à paramétrer simultanément les commandes et les variables d'état. L'algorithme est le suivant :

1. Paramétrer les commandes et les variables d'état en un nombre fini de variables de décision,
2. Discrétiser les équations différentielles (4.2) aux instants sélectionnés. Les étapes 1 et 2 transforment le problème d'optimisation (4.1)-(4.3) en un problème de Programmation Non-Linéaire (PNL),
3. Choisir une initialisation des variables de décision,
4. Utiliser un algorithme pour résoudre le nouveau problème d'optimisation (par exemple la Programmation Quadratique Séquentielle, dite SQP).

Généralement, les trajectoires des variables d'état sont approchées par des fonctions polynomiales. Seuls les coefficients de ces polynômes sont optimisés, ce qui a pour résultat de réduire considérablement le nombre de variables d'optimisation. De plus, le problème d'optimisation obtenu après paramétrisation de la commande et des états possède une structure creuse, ce qui réduit les temps de calcul et rend le problème plus robuste vis-à-vis de l'initialisation [Tebbani *et al.*, 2006].

Quelques travaux présentent les différentes applications de cette approche. Dans [Renfro *et al.*, 1987] par exemple, elle est utilisée pour déterminer le profil optimal de la température afin de maximiser la production d'un produit intermédiaire. Dans [Tebbani *et al.*, 2006], elle est utilisée pour optimiser le fonctionnement du bioprocédé de Park et Ramirez et celui de la culture de pénicilline.

## 4.4 Résolution par l'approche indirecte

Dans l'approche dite directe, la commande optimale est approchée par une commande en dimension finie. Dans ce paragraphe, on recherche plutôt la solution optimale sous forme de fonction continue.

### 4.4.1 Formulation du problème d'optimisation

En utilisant le Principe du Maximum, le problème d'optimisation dynamique (4.1)-(4.3) est reformulé comme suit, en introduisant la fonction Hamiltonienne  $\mathcal{H}(t)$  [Bryson et Ho, 1975] :

$$\min_{u(t), \mu(t), v} \mathcal{H} = L(x, u) + \lambda^T \mathcal{F}(x, u) + \mu^T \psi(x, u) \quad (4.4)$$

sous les contraintes :

$$\dot{x} = \mathcal{F}(x, u), \quad x(0) = x_0 \quad (4.5)$$

$$\dot{\lambda}^T = -\frac{\partial \mathcal{H}}{\partial x} \quad (4.6)$$

$$\lambda^T(t_f) = \frac{\partial \phi}{\partial x} \Big|_{t_f} + v^T \left( \frac{\partial \varphi}{\partial x} \right) \Big|_{t_f} \quad (4.7)$$

où :

- $\lambda(t) \neq 0$  est le vecteur des états adjoints, de dimension  $n$ ,

- $\mu(t) \geq 0$  est le vecteur des multiplicateurs de Karush-Kuhn-Tucker (KKT) associés aux contraintes sur la trajectoire  $\psi$ , de dimension  $n_\psi$ ,
- $\nu \geq 0$  est le vecteur des multiplicateurs de KKT pour les contraintes terminales  $\varphi$ , de dimension  $n_\varphi$ .

Les multiplicateurs de KKT  $\mu$  et  $\nu$  sont strictement positifs quand les contraintes correspondantes sont actives et nuls sinon. Ainsi, la propriété  $\mu^T \psi(x, u) = 0$ ,  $\nu^T \varphi(x(t_f)) = 0$  est toujours vraie.

Il est à noter que dans le cas où la contrainte terminale est une contrainte d'égalité (soit de la forme  $\varphi(x(t_f)) = 0$ ), les variables  $\nu$  sont les multiplicateurs de Lagrange. Ils sont dans ce cas des réels, sans aucune condition sur leurs signes.

La contrainte (4.7) caractérise la valeur des états adjoint à l'instant final. Cette contrainte est connue sous le nom de contrainte de transversalité.

La commande optimale est la commande qui minimise le Hamiltonien, soit, en appliquant la condition nécessaire d'optimalité :

$$\mathcal{H}_u(u^*) = 0 \quad (4.8)$$

Cette condition est généralisée dans le cas des systèmes avec commande saturée par le principe du Maximum de Pontryagin. Ce cas particulier sera détaillé à la section 4.6 puisque le système étudié fait partie de cette classe particulière de systèmes.

En résumé, le problème d'optimisation est équivalent à résoudre le système d'équations suivant :

$$\dot{x} = \mathcal{F}(x, u), \quad x(0) = x_0 \quad (4.9.a)$$

$$\dot{\lambda}^T = -\frac{\partial \mathcal{H}}{\partial x} = -\frac{\partial L}{\partial x} - \lambda^T \frac{\partial \mathcal{F}}{\partial x} - \mu^T \frac{\partial \psi}{\partial x} \quad (4.9.b)$$

$$\lambda^T(t_f) = \frac{\partial \phi}{\partial x} \Big|_{t_f} + \nu^T \left( \frac{\partial \varphi}{\partial x} \right) \Big|_{t_f} \quad (4.9.c)$$

$$\mu^T \psi(x, u) = 0 \quad (4.9.d)$$

$$\nu^T \varphi(x(t_f)) = 0 \quad (4.9.e)$$

$$\frac{\partial \mathcal{H}}{\partial u} = 0 \quad (4.9.f)$$

## 4.4.2 Résolution

Afin de résoudre le système d'équations (4.9), trois méthodes différentes existent dans la littérature :

- méthode de tir,
- méthode du gradient,
- paramétrisation des variables d'états et des états adjoints.

#### 4.4.2.1 Méthode de tir

Cette méthode consiste à déterminer les fonctions  $x$ ,  $u$ ,  $\lambda$ ,  $\mu$  et  $\nu$  satisfaisant les conditions (4.9). Les équations (4.9.a) et (4.9.b) sont résolues simultanément par intégration des deux dynamiques.

Le système d'équations (4.9) est un problème aux deux bouts (Two-Point Boundary Value Problem, TPBVP). En effet, les valeurs des variables d'état et des états adjoints sont connues, respectivement à l'instant initial et à l'instant final. Or, les valeurs des états adjoints à l'instant initial et les valeurs des états à l'instant final sont inconnues. Pour résoudre ce type de problème, une technique classique consiste en la méthode de tir. Cette méthode recherche la solution du problème (4.9) par essais successifs selon l'algorithme suivant :

1. Paramétrer  $\mu$  en utilisant un nombre fini de variables. Le vecteur des variables de décision inclut également  $\lambda(0)$  et  $\nu$ ,
2. Choisir une initialisation des variables de décision,
3. Résoudre (4.9.f) analytiquement pour  $u(x, \lambda)$ ,
4. Intégrer les équations différentielles (4.9.a) et (4.9.b) en utilisant  $x(0)$ ,  $\lambda(0)$  et  $u(x, \lambda)$  obtenues à partir de l'étape 3. En déduire la valeur des états adjoints à l'instant final,  $\lambda(t_f)$ ,
5. Vérifier si les équations (4.9.a) et (4.9.b) sont satisfaites avec la nouvelle valeur de  $\lambda(t_f)$ . Mettre à jour les variables de décision. Répéter les étapes 4 et 5 jusqu'à convergence.

Quelques applications de cette méthode sont présentées dans la littérature, notamment pour la polymérisation [Hicks *et al.*, 1969] et [Sacks *et al.*, 1972]. Cependant, cette méthode possède des inconvénients, à savoir :

- la forte sensibilité de l'algorithme liée à l'initialisation de la valeur initiale des états adjoints et donc la difficulté de convergence pour les méthodes de résolution itérative,
- la nécessité d'utiliser des processus d'initialisation aléatoire pour les algorithmes d'optimisation, d'où une augmentation importante du temps de calcul,
- la difficulté pour déterminer les instants de commutation et les arcs singuliers (cette terminologie sera définie par la suite).

#### 4.4.2.2 Méthode du gradient

La méthode du gradient part du même principe que l'approche séquentielle, sauf que dans ce cas, le gradient est calculé à partir de l'équation (4.9.f).

L'algorithme de la méthode du gradient est donné ci-dessous :

1. Paramétrer  $u$  et  $\mu$  en utilisant un nombre fini de variables. Le vecteur des variables de décision inclut également  $\nu$ ,
2. Choisir une initialisation des variables de décision,
3. Intégrer les équations différentielles (4.9.a) de 0 à  $t_f$ ,
4. Intégrer les équations des états adjoints (4.9.b) de  $t_f$  à 0. Calculer le gradient  $\partial \mathcal{H} / \partial u$  en utilisant (4.9.f),
5. Utiliser un algorithme d'optimisation pour mettre à jour les valeurs des variables de décision. Répéter les étapes 3-5 jusqu'à minimiser  $\mathcal{H}$ .

Cette méthode est largement appliquée dans la littérature. Elle est utilisée pour l'optimisation du fonctionnement d'un procédé chimique dans [Jaspan et Coull, 1972] et de colonnes à distiller dans [Diwekar, 1995]. L'initialisation des variables de décision n'a pas d'influence sur la convergence. Néanmoins, le nombre de variables de décision à paramétrer est important et la convergence est lente près de l'optimum.

#### 4.4.2.3 Paramétrisation des états et des états adjoints

Cette méthode consiste à paramétrer les variables d'état et les états adjoints. Deux techniques de résolution numérique sont présentées dans ce cas :

- résolution par équations algébriques non-linéaires,
- résolution par quasi-linéarisation.

Les algorithmes de ces deux techniques sont présentés dans ce qui suit.

##### 4.4.2.3.1 Résolution par équations algébriques non-linéaires

Cette approche est similaire à l'approche simultanée présentée précédemment (paragraphe 4.3.2). L'algorithme de cette technique est le suivant :

1. Paramétrer  $x$ ,  $u$ ,  $\lambda$  et  $\mu$  en utilisant un nombre fini de variables de décision. Le vecteur des variables de décision inclut  $\nu$ ,
2. Discrétiser les équations différentielles (4.9.a)-(4.9.b) et la condition (4.9.f) pour un nombre fini d'instant  $t$ . Ces deux étapes transforment le problème de résolution des équations différentielles algébriques non-linéaires en un problème de résolution d'équations algébriques non-linéaires,
3. Choisir une initialisation des variables de décision,
4. Résoudre le nouveau problème d'optimisation.

Cette approche a été développée dans [Goh et Teo, 1988] pour résoudre des problèmes d'optimisation avec contraintes. L'étape 4 de cette technique est résolue en utilisant l'algorithme de Newton-Raphson dans [Schwarz, 1989].

#### 4.4.2.3.2 Solution par quasi-linéarisation

Cette approche consiste à utiliser la méthode des deux bouts avec des linéarisations successives. L'algorithme est le suivant :

1. Paramétrer  $x$ ,  $\lambda$  et  $\mu$  en utilisant un nombre fini de variables de décision. Le vecteur des variables de décision inclut  $v$ ,
2. Choisir une initialisation des variables de décision,
3. Linéariser les équations différentielles (4.9.a) et (4.9.b), ce qui transforme le problème de résolution d'équations différentielles algébriques non-linéaires en un problème de résolution d'équations différentielles algébriques linéaires,
4. Résoudre le nouveau problème d'optimisation,
5. Réinitialiser les variables de décision avec la solution de l'étape 4 et répéter les étapes 3-5 jusqu'à convergence.

Son développement théorique est présenté dans [Bryson et Ho, 1975]. Cette approche a été appliquée dans le domaine des procédés chimiques dans [Lee, 1968].

## 4.5 Récapitulatif des méthodes de résolution

Le Tableau 4.1 représente la classification des méthodes présentées précédemment.

Nature des variables d'état et des commandes	Formulation du problème d'optimisation	
	Directe	Indirecte
Variables d'état continues Commandes continues	-	Méthode de tir
Variables d'état continues Commandes paramétrées	Approche séquentielle	Méthode du gradient
Variables d'état paramétrées Commandes paramétrées	Approche simultanée	Discrétisation des variables d'état et des états adjoints

**Tableau 4.1.** Classification des méthodes de résolution numérique d'un problème d'optimisation dynamique.

Ainsi, deux grandes familles d'algorithmes pour la résolution de problèmes d'optimisation émergent de cette classification :

- Les méthodes de type approche directe présentent l'avantage d'être simple à mettre en œuvre puisqu'une approximation de la solution optimale peut être aisément calculée. La difficulté lors de l'application de ces méthodes est leur convergence vers des

minimas locaux. En effet, du fait de la paramétrisation du problème, elles convergent généralement vers des solutions sous-optimales dont la qualité dépend principalement de la paramétrisation choisie et surtout de l'initialisation. Ces méthodes sont donc principalement utilisées pour rechercher des solutions faisables, ou des solutions sous-optimales pouvant servir d'initialisation à une résolution par approche indirecte.

- Les méthodes de type approche indirecte présentent l'avantage de converger vers la solution optimale continue. Cependant, elles sont très sensibles à l'initialisation des valeurs des états adjoints à l'instant initial et à la détermination des instants de commutation est numériquement très difficile. Ainsi, l'application de ces méthodes est très délicate.

L'application de ces méthodes est d'autant plus difficile dans notre cas que le système possède un modèle avec commutation et que la commande optimale est de type commande singulière comme nous le verrons plus loin.

Pour remédier à ces difficultés, le choix final a été de reprendre les développements proposés par [Srinivasan *et al.*, 2003] qui consistent à appliquer la formulation du principe du Maximum de Pontryagin non pas pour résoudre le problème d'optimisation mais pour caractériser la solution optimale. Cette caractérisation facilite ensuite l'application d'une méthode directe de résolution numérique.

## 4.6 Principe du Maximum de Pontryagin

Dans le cas du bioprocédé étudié, la détermination d'un profil optimal d'alimentation revient à résoudre un problème d'optimisation d'une fonction coût à l'instant final avec un modèle affine en la commande. En effet, les modèles de bioprocédés sont des fonctions linéaires en débit d'alimentation. De plus, la commande du système est généralement saturée :

$$0 \leq u \leq F_{\max} \quad (4.10)$$

Le principe du Maximum présenté à la section 4.4 se généralise et devient le principe du Maximum de Pontryagin. Cette partie se penche sur les détails de ce principe.

Soit le problème d'optimisation suivant (seul un coût terminal est considéré pour notre cas) :

$$\min_{u(t)} J = \min_{u(t)} \phi(x(t_f)) \quad (4.11.a)$$

$$\text{sous les contraintes : } \dot{x} = \mathcal{F}(x, u) = f(x) + g(x)u, \quad x(0) = x_0 \quad (4.11.b)$$

$$\psi(x, u) = \psi^f(x) + \psi^g(x)u \leq 0, \quad \phi(x(t_f)) \leq 0 \quad (4.11.c)$$

$$u_{\min} \leq u \leq u_{\max} \quad (4.11.d)$$

où :

- $J$  est le critère (fonction objectif) à minimiser,
- $x$  est le vecteur des variables d'état, de dimension  $n$ ,
- $x_0$  est le vecteur initial des variables d'état et est connu,
- $u$  est le vecteur des commandes, de dimension  $m = 1$ ,

- $\psi$  est le vecteur des contraintes sur la trajectoire (incluant les contraintes sur les variables d'état et sur les commandes), de dimension  $n_\psi$ ,
- $\varphi$  est le vecteur des contraintes terminales, de dimension  $n_\varphi$ ,
- $\mathcal{F}$  est une fonction décrivant la dynamique du système,
- $\phi$  est une fonction scalaire représentant le coût terminal,
- $t_f$  est le temps final, supposé fixé.

Ce problème présente la particularité d'être affine en la commande (dans la dynamique et dans les contraintes). Une autre caractéristique importante est que la commande est saturée.

L'application du principe du Maximum de Pontryagin à la résolution du problème d'optimisation dynamique (4.11) permet de définir la fonction Hamiltonienne  $\mathcal{H}$  suivante [Bryson et Ho, 1975] :

$$\mathcal{H} = \lambda^T f(x) + \mu^T \psi^f(x) + (\lambda^T g(x) + \mu^T \psi^g(x))u \quad (4.12)$$

où :

- $\lambda$  est le vecteur des états adjoints,
- $\mu$  représente les multiplicateurs de KKT.

Les équations (4.9) sont également obtenues. La seule particularité du principe du Maximum réside dans le calcul de la commande optimale. En effet, la condition nécessaire  $\partial \mathcal{H} / \partial u = 0$  donne :

$$\mathcal{H}_u = \lambda^T g(x) + \mu^T \psi^g(x) = 0 \quad (4.13)$$

On constate que  $\mathcal{H}_u$  ne dépend pas de  $u$ . Ainsi, l'expression analytique de la commande optimale ne peut pas être obtenue directement par (4.13).

Le principe de Pontryagin caractérise le calcul de la commande optimale  $u^*$  par :

$$u^* = \underset{u \in [u_{\min}, u_{\max}]}{\text{Arg min}} \mathcal{H} \quad (4.14)$$

Par ailleurs, la commande est bornée (cf. (4.11.d)). Dans ce cas, l'application du principe du Maximum de Pontryagin donne alors trois possibilités pour la commande optimale :

$$u^* = \begin{cases} u_{\max} & \text{si } \mathcal{H}_u > 0 \\ u_{\min} & \text{si } \mathcal{H}_u < 0 \\ \text{Arc singulier} & \text{si } \mathcal{H}_u = 0 \end{cases} \quad (4.15)$$

Et d'après les conditions de Karush Kuhn Tucker (KKT), la commande optimale est obtenue comme suit :

- Soit  $\mu(t) \neq 0$ , dans ce cas la contrainte correspondante est active ( $\psi = 0$ ) et la commande optimale (que l'on notera  $u_{path}$ ) est donc définie afin de suivre cette contrainte,

- Soit  $\mu(t) = 0$ , dans ce cas la contrainte correspondante n'est pas active ( $\psi \neq 0$ ) et la commande optimale est soit égale aux valeurs limites, soit obtenue en annulant les dérivées temporelles successives de  $\mathcal{H}_u$  jusqu'à ce que la commande apparaisse explicitement.

## 4.7 Caractérisation analytique de la solution optimale

La solution optimale recherchée peut être de deux natures :

- soit la trajectoire optimale se situe sur la frontière du domaine de faisabilité, la commande est alors non singulière. C'est la commande notée  $u_{path}$  qui maintient le système sur la contrainte définissant cette frontière,
- soit elle ne se trouve pas sur une contrainte, auquel cas la commande est soit égale à ses valeurs limites, soit singulière (selon (4.15)).

Ces différentes lois de commande peuvent coexister dans un problème d'optimisation avec des instants de commutation entre les différentes zones.

### 4.7.1 Caractérisation de la commande non singulière

Un intervalle non singulier est caractérisé par au moins une des contraintes sur la trajectoire qui est active. Les éléments correspondants dans le multiplicateur  $\mu$  sont alors non nuls.

Dans ce cas,  $\mathcal{H}_u$  est dérivé jusqu'à ce qu'un multiplicateur  $\mu$  non nul satisfasse la relation  $d^j \mathcal{H}_u / dt^j = 0$ , ou que le terme  $\lambda^T \Delta^j g$  soit non nul, où l'opérateur  $\Delta$  est donné par l'expression suivante :

$$\Delta h = \left( \frac{\partial h}{\partial x} f - \frac{\partial f}{\partial x} h \right) + u \left( \frac{\partial h}{\partial x} g - \frac{\partial g}{\partial x} h \right) \quad (4.16)$$

Cet opérateur représente la différentiation d'une fonction  $h$  le long de la trajectoire du système (4.11.b).

Généralement, une seule des contraintes sera active (la contrainte la plus forte), le multiplicateur  $\mu$  associé indiquera alors la contrainte à partir de laquelle la commande  $u$  peut être déterminée. Ainsi, la commande est déterminée en dérivant plusieurs fois la contrainte, jusqu'à ce que la commande apparaisse explicitement (pour la contrainte active).

Dans le cas général, le calcul de la commande optimale dépend de l'ordre de dérivation  $j$  :

- Si  $j = 0$ , une fonction  $\mu$  doit satisfaire la relation  $\mathcal{H}_u = 0$ . De  $\mu^T (\psi^f + \psi^g u) = 0$  et de  $\mu^T \psi^g \neq 0$ , la commande optimale est alors donnée par :

$$u_{path} = -\mu^T (\psi^f + \psi^g u) / \mu^T \psi^g \quad (4.17)$$

- Si  $j > 0$  puisque la commande n'apparaît pas explicitement avant la  $j^{\text{ème}}$  dérivation, la contrainte active a seulement une dépendance implicite vis-à-vis de la commande. La solution optimale est dans ce cas donnée par :

$$u_{path} = -\mu^T \frac{\partial}{\partial x} \frac{d^{j-1}\psi}{dt^{j-1}} (f + gu) \Big/ \mu^T \frac{\partial \psi}{\partial x} \Delta^{j-1} g \quad (4.18)$$

## 4.7.2 Caractérisation de la commande singulière

Les paragraphes ci-dessous détaillent l'expression de la commande singulière et déterminent les intervalles singuliers.

### 4.7.2.1 Calcul analytique de la commande singulière

Lorsque la contrainte n'est pas active (soit  $\mu(t) = 0$ ), la dérivée première du Hamiltonien par rapport à  $u$  se simplifie et devient :

$$\mathcal{H}_u = \lambda^T g(x) \quad (4.19)$$

Par ailleurs, la commande singulière est obtenue en annulant les dérivées temporelles successives de  $\mathcal{H}_u$  :

$$\frac{d\mathcal{H}_u}{dt} = 0 \quad \text{soit} \quad \lambda^T q(x) = 0 \quad (4.20)$$

avec :

$$q(x) = \frac{\partial g(x)}{\partial x} f(x) - \frac{\partial f(x)}{\partial x} g(x) \quad (4.21)$$

Le résultat du calcul de la première dérivée de la fonction  $\mathcal{H}_u$  (4.20) montre que la commande  $u$  n'apparaît pas explicitement, le calcul de la dérivée seconde de  $\mathcal{H}_u$  est alors indispensable :

$$\frac{d^2\mathcal{H}_u}{dt^2} = 0 \quad \text{soit} \quad -\lambda^T \frac{\partial f}{\partial x} q(x) - \lambda^T \frac{\partial g}{\partial x} q(x)u + \lambda^T \frac{\partial q(x)}{\partial x} (f(x) + g(x)u) = 0 \quad (4.22)$$

La dérivée seconde de la fonction  $\mathcal{H}_u$  fait apparaître la commande  $u$  (4.22).

La commande singulière  $u_{\text{sing}}$  s'écrit alors sous la forme :

$$u_{\text{sing}} = \frac{\lambda^T \left( \frac{\partial f}{\partial x} q - \frac{\partial q}{\partial x} f \right)}{\lambda^T \left( \frac{\partial q}{\partial x} g - \frac{\partial g}{\partial x} q \right)} \quad \text{si :} \quad \lambda^T \left( \frac{\partial q}{\partial x} g - \frac{\partial g}{\partial x} q \right) \neq 0 \quad (4.23)$$

L'expression explicite de la commande singulière  $u_{\text{sing}}$  est difficile à déterminer à partir de la formule (4.23) puisque les états adjoints  $\lambda$  sont inconnus. Pour y remédier, une approche de caractérisation, basée sur l'étude des intervalles singuliers, est présentée dans ce qui suit.

### 4.7.2.2 Détermination des intervalles singuliers

Sur un intervalle singulier, aucune des contraintes n'est active. Par conséquent, les dérivées successives de (4.19) mènent à :

$$\frac{d^j \mathcal{H}_u}{dt^j} = \lambda^T \Delta^j g = 0, \quad \forall j \quad (4.24)$$

où l'opérateur  $\Delta$  est défini par (4.16).

L'équation (4.24) peut s'écrire sous la forme matricielle suivante :

$$\lambda^T \begin{bmatrix} \Delta^0 g & \Delta^1 g & \dots & \Delta^{n-1} g \end{bmatrix} \equiv \lambda^T M = 0 \quad (4.25)$$

Puisque les états adjoints  $\lambda$  sont non nuls, la formule (4.25) indique que, pour une combinaison spécifique de l'état et de la commande, la singularité se produit quand la matrice  $M$  n'est plus de rang plein. Ainsi, on observe une dégénérescence de rang de la matrice  $M$  le long de la singularité. Cette remarque permet de caractériser facilement la présence d'arcs singuliers.

La nature de la loi de commande optimale dépend de l'ordre de la singularité, noté  $\sigma$ , qui est le nombre minimal des dérivées temporelles de  $\mathcal{H}_u$  nécessaires pour que la commande apparaisse explicitement (i.e. quand le terme  $\Delta^j g$  dépend explicitement de la commande  $u$ ).

Selon la valeur de  $\sigma$ , la classification suivante des lois singulières peut être faite :

- si  $\sigma + 1 < \text{rang}(M)$  : la loi optimale est dynamique,
- si  $\sigma + 1 = \text{rang}(M)$  : la loi optimale est statique,
- si  $\text{rang}(M) < \sigma + 1 < \infty$  : la loi optimale correspond à une surface,
- si  $\sigma = \infty$  : le rang de  $M$  est indépendant des états et de la commande. Deux cas de figures sont à distinguer :
  - soit  $\text{rang}(M) = n$ , alors il n'existe pas de surface singulière,
  - soit  $\text{rang}(M) < n$ , selon la fonction de coût, soit il n'existe pas de commande singulière, soit la commande singulière n'est pas unique.

Ainsi, pour connaître la nature de la commande optimale singulière à appliquer au système, il convient de calculer le rang de la matrice  $M$  et de déterminer la commande associée. Dans la section suivante, cette stratégie sera appliquée au cas de commande du bioprocédé de culture des *E. coli*. La présence de la commande singulière sera démontrée et sa nature (statique, dynamique ou surface) sera mise en évidence, en étudiant le rang de la matrice  $M$ .

## 4.8 Application à la détermination d'un profil optimal de commande du bioprocédé *E. coli*

La détermination du profil optimal de commande de la culture des *E. coli* a fait l'objet d'études dans la littérature. Dans [Rocha, 2004], une approche directe par paramétrisation de la commande et utilisation d'algorithmes génétiques est utilisée pour la maximisation de la croissance de la biomasse. Le résultat obtenu est un profil optimal exponentiel, mais purement numérique.

Dans [Ko et Wang, 2006], la maximisation de la production de protéines est considérée par une technique de type « run-tu-run optimization ». Elle consiste à alterner des phases d'identification et d'optimisation pour affiner la recherche de la solution optimale.

Cette partie présente l'application de la théorie développée au paragraphe 4.7 au fonctionnement du bioprocédé de cultures *E. coli*. Pour faciliter cette application, une première étape consiste à simplifier le modèle de commande présenté au chapitre 2.

#### 4.8.1 Mise en œuvre d'un modèle simplifié du bioprocédé *E. coli*

Rappelons que le modèle dynamique du bioprocédé présenté au chapitre 2 est :

$$\begin{aligned}\frac{dX}{dt} &= r_X X - \frac{F_{in}}{W} X \\ \frac{dS}{dt} &= -q_S X + \frac{F_{in}}{W} (S_{in} - S) \\ \frac{dA}{dt} &= r_A X - \frac{F_{in}}{W} A \\ \frac{dW}{dt} &= F_{in}\end{aligned}\tag{4.26}$$

où :

- $X$ ,  $S$  et  $A$  sont respectivement les concentrations en biomasse, en substrat (glucose) et en acétate,
- $W$  est le poids,
- $r_X$ ,  $r_A$  et  $q_S$  sont les taux spécifiques dépendants des variables d'état et décrits au chapitre 2,
- $F_{in}$  est le débit d'alimentation,
- $S_{in}$  représente la concentration en substrat dans le milieu d'alimentation.

La réécriture du modèle dynamique (4.26) est effectuée en considérant la nouvelle représentation suivante pour les variables d'état :

$$(XW, SW, AW, W)\tag{4.27}$$

Avec ce nouveau jeu de variables, la dynamique du modèle de culture des *E. coli* devient :

$$\begin{aligned}\frac{d}{dt}(XW) &= r_X XW \\ \frac{d}{dt}(SW) &= -q_S XW + S_{in}u \\ \frac{d}{dt}(AW) &= r_A XW \\ \frac{dW}{dt} &= F_{in}\end{aligned}\tag{4.28}$$

où le débit d'alimentation  $F_{in}$  représente la commande  $u$  que l'on cherche à optimiser.

Le modèle (4.28) est toujours de dimension 4. Un autre changement de variable peut s'envisager pour aller encore plus loin. Ce dernier est basé sur la relation entre les variables d'état. En effet, il est possible d'exprimer la variable  $SW$  qui, correspond au substrat, en fonction des autres variables d'état. Un nouveau changement de variable, consistant à remplacer  $SW$  par des fonctions des trois variables restantes, permet donc de réduire la taille du vecteur d'état. Dans ce qui suit, la relation entre  $SW$  et les autres variables du modèle (4.28) est donnée, selon les régimes de fonctionnement :

- **Pour le régime oxydatif ( $\mu_2 = 0$ ) :**

Rappelons que pour ce régime de fonctionnement, deux sous-régimes existent : sous-régime 1 et sous-régime 2.

Pour le sous-régime 1, les expressions des taux  $r_X$  et  $r_A$  sont :

$$\begin{aligned} r_X &= \frac{q_S}{k_1} + \frac{q_{AC}}{k_4} \\ r_A &= -q_{AC} \end{aligned} \quad (4.29)$$

Pour le sous-régime 2, les expressions des taux  $r_X$  et  $r_A$  s'écrivent :

$$\begin{aligned} r_X &= \frac{q_S}{k_1} + \frac{k_{os}}{k_4} \frac{q_{Scrit} - q_S}{k_{oa}} \\ r_A &= -k_{os} \frac{q_{Scrit} - q_S}{k_{oa}} \end{aligned} \quad (4.30)$$

Dans les deux cas (4.29) et (4.30), on peut écrire :

$$r_X = \frac{q_S}{k_1} - \frac{r_A}{k_4} \quad (4.31)$$

A partir des quatre équations différentielles décrivant le modèle (4.28) et de la relation (4.31), la dérivée de  $SW$  s'écrit sous la forme suivante :

$$\begin{aligned} \frac{d(SW)}{dt} &= -k_1 \frac{d(XW)}{dt} - \frac{k_1}{k_4} \frac{d(AW)}{dt} + S_{in} \frac{dW}{dt} \\ \text{soit } \frac{d}{dt} \left( SW + k_1 XW + \frac{k_1}{k_4} AW - S_{in} W \right) &= 0 \end{aligned} \quad (4.32)$$

La concentration en substrat  $S$  peut donc être exprimée en fonction des autres variables d'état et en fonction des conditions initiales par la formule :

$$S = \frac{1}{W} \left( I_0 - k_1 XW - \frac{k_1}{k_4} AW \right) + S_{in} \quad (4.33)$$

avec :

$$I_0 = S_0 W_0 + k_1 X_0 W_0 + \frac{k_1}{k_4} A_0 W_0 - S_{in} W_0 \quad (4.34)$$

- **Pour le régime oxydo-fermentatif ( $\mu_3 = 0$ ) :**

Les expressions des taux  $r_X$  et  $r_A$  pour ce régime de fonctionnement sont :

$$\begin{aligned} r_X &= \frac{q_S - q_{S \text{ crit}}}{k_2} + \frac{q_{S \text{ crit}}}{k_1} \\ r_A &= k_3 \frac{q_S - q_{S \text{ crit}}}{k_2} \end{aligned} \quad (4.35)$$

De la même manière que dans le cas du régime oxydatif, on exprime la dérivée de  $SW$  en fonction des autres variables du modèle (4.28) pour en déduire une expression de la concentration en substrat en fonction des autres variables :

$$\begin{aligned} \frac{d(SW)}{dt} &= -k_1 \frac{d(XW)}{dt} + k_1 \beta \frac{d(AW)}{dt} + S_{in} \frac{dW}{dt} \\ \text{soit } \frac{d}{dt}(SW + k_1 XW - k_1 \beta AW - S_{in} W) &= 0 \end{aligned} \quad (4.36)$$

avec :  $\beta = (1/k_3)(1 - k_2/k_1)$ .

L'expression de la concentration en substrat  $S$  est alors donnée par :

$$S = \frac{1}{W} (J_0 - k_1 XW + k_1 \beta AW) + S_{in} \quad (4.37)$$

$$\text{avec : } J_0 = (S_0 W_0 + k_1 X_0 W_0 - k_1 \beta A_0 W_0 - S_{in} W_0) \quad (4.38)$$

Il est donc possible de retirer la quantité de substrat du vecteur d'état et ne garder que les trois variables  $XW$ ,  $AW$  et  $W$  représentant respectivement les quantités de biomasse, d'acétate et le poids. Le nouveau système est alors donné par :

$$\frac{d}{dt} \begin{pmatrix} XW \\ AW \\ W \end{pmatrix} = \begin{pmatrix} r_X XW \\ r_A XW \\ F_{in} \end{pmatrix} \quad (4.39)$$

La concentration en substrat  $S$  est donnée par (4.33) ou (4.37) selon le type de régime.

Ainsi, avec cette nouvelle représentation du système, la dynamique du système est alors donnée par :

$$\dot{x} = f(x) + g(x)u \quad (4.40)$$

$$\text{avec : } x = \begin{bmatrix} XW \\ AW \\ W \end{bmatrix}, \quad f = \begin{bmatrix} r_X XW \\ r_A XW \\ 0 \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad u = F_{in}.$$

Par la suite, seule cette représentation réduite sera considérée.

## 4.8.2 Définition du problème d'optimisation

Le problème d'optimisation considéré consiste à maximiser la croissance de la biomasse, dans une culture fed-batch, en optimisant le débit d'alimentation.

En effet, la culture des bactéries *E. coli* a pour finalité dans notre application de fabriquer des médicaments ou des vaccins. Ainsi, l'objectif est de déterminer le profil temporel de débit

d'alimentation  $u(t)$  qui maximise la quantité de biomasse à l'instant final  $X(t_f)W(t_f)$ . Par ailleurs, maximiser le terme  $X(t_f)W(t_f)$  revient à minimiser  $-X(t_f)W(t_f)$ .

La fonction coût à minimiser est donc :

$$\phi(x(t_f)) = -X(t_f)W(t_f) \quad (4.41)$$

La commande qui est le débit d'alimentation est saturée, puisqu'elle est limitée par les capacités des pompes utilisées.

Pour ce qui est des contraintes sur les états, la trajectoire est régie par la dynamique (4.40) et les états sont des réels positifs (puisque les états sont des concentrations et le poids du bioprocédé). Une limitation sur le poids peut également être imposée. Enfin, l'existence de deux régimes de fonctionnement (oxydatif ou oxydo-fermentatif) sera gérée à l'aide d'une contrainte de la forme  $\psi(x(t_1), u(t_1)) = 0$ , où  $t_1$  est l'instant de commutation d'un régime de fonctionnement à l'autre et qui devra être optimisé.

Le problème d'optimisation est donc le suivant :

$$\begin{aligned} & \min_{u(t)} [-X(t_f)W(t_f)] \\ & \text{sous les contraintes :} \quad \begin{aligned} \dot{x} &= f(x) + g(x)u, & x(0) &= x_0 \\ q_S(t_1) &= q_{Scrit}(t_1), & t_1 &\in [0, t_f] \\ 0 &\leq u(t) \leq F_{\max} \\ 0 &\leq W(t) \leq W_{\max} \\ 0 &\leq x(t) \end{aligned} \end{aligned} \quad (4.42)$$

Les solutions optimales possibles sont présentées dans ce qui suit en appliquant la théorie développée au paragraphe 4.7.

### 4.8.3 Expressions analytiques des solutions optimales

Il a été établi au paragraphe 4.7 que la solution optimale peut être de deux natures :

- soit la contrainte est active, dans ce cas la commande optimale permet de suivre la frontière, soit  $u_{opt} = u_{path}$ ,
- soit la contrainte n'est pas active, dans ce cas trois solutions sont possibles :
  - la commande est saturée à sa valeur minimale :  $u_{opt} = 0$ ,
  - la commande est saturée à sa valeur maximale :  $u_{opt} = F_{\max}$ ,
  - la commande est singulière :  $u_{opt} = u_{sing}$

La solution optimale peut aussi être composée de ces différentes lois de commande avec des instants de commutation qu'il faut également optimiser. La solution optimale recherchée dépend du critère, des conditions initiales et finales et du temps total  $t_f$ . Ainsi, selon le cas d'étude considéré, le scénario optimal peut différer et sera forcément composé d'un ou plusieurs arcs optimaux, correspondant aux différentes solutions optimales listées ci-dessus.

#### 4.8.3.1 Expression analytique de la commande non singulière $u_{path}$

La détermination de l'expression de la commande  $u_{path}$  permettant le maintien de la trajectoire sur la frontière revient à trouver la commande qui maintient le système à la frontière des deux régimes de fonctionnement : les régimes oxydatif et oxydo-fermentatif.

La contrainte active est ici la condition à la frontière des deux régimes et s'écrit sous la forme :

$$q_S = q_{Scrit} \quad (4.43.a)$$

$$\text{où :} \quad q_S = q_{S \max} \frac{S}{k_s + S} \quad \text{et} \quad q_{Scrit} = \frac{q_{O \max}}{k_{os}} \frac{k_{io}}{k_{io} + A} \quad (4.43.b)$$

En appliquant les formules détaillées au paragraphe 4.7.1, il vient que la commande est caractérisée par :

$$\frac{dq_S}{dt} = \frac{dq_{Scrit}}{dt} \quad (4.43.c)$$

Ainsi, une seule dérivation suffit (i.e.  $j=1$ ) et l'expression de la commande non singulière  $u_{path}$  est alors donnée à partir de (4.43.c) et des équations différentielles du modèle (4.28) :

$$u_{path} = \frac{\left( r_A \frac{\partial q_{scrit}}{\partial A} + q_s \frac{\partial q_s}{\partial S} \right)}{\left( A \frac{\partial q_{scrit}}{\partial A} + (S_{in} - S) \frac{\partial q_s}{\partial S} \right)} XW \quad (4.44.a)$$

$$\text{où :} \quad \frac{\partial q_{scrit}}{\partial A} = \frac{-q_{scrit}}{k_{io} + A} \quad \text{et} \quad \frac{\partial q_s}{\partial S} = q_s \left( \frac{1}{S} - \frac{1}{k_s + S} \right) \quad (4.44.b)$$

Ainsi, contrairement à la commande singulière, la commande  $u_{path}$  ne dépend pas du régime de fonctionnement.

#### Remarque :

Dans le cas particulier où la concentration en acétate est nulle (soit  $A = 0$ ), l'expression (4.44.a) se simplifie et devient :

$$u_{path} \Big|_{A=0} = \frac{q_s}{S_{in} - S} XW \quad (4.44.c)$$

Or, d'après (4.43.a), dans ce cas, la concentration en substrat est également constante, soit  $S = S^* \equiv \frac{q_{O \max} k_s}{q_{S \max} k_{os} - q_{O \max}}$ . Ainsi, la biomasse suivant une loi exponentielle par intégration de l'équation différentielle (4.39), la commande optimale est donnée par :

$$u_{path} \Big|_{A=0} = \frac{q_{O \max} / k_{os}}{S_{in} - S^*} X_0 W_0 e^{\frac{q_{O \max}}{k_{os} k_1} (t - t_0)} \quad (4.44.d)$$

### 4.8.3.2 Expression analytique de la commande singulière $u_{sing}$

Comme démontré précédemment, la présence ou l'absence d'une commande singulière est déterminée par l'étude de la matrice  $M$  définie dans (4.25).

Avec la représentation donnée en (4.40), la matrice  $M$  est une matrice de taille  $3 \times 3$  :

$$M = [\Delta^0 g \quad \Delta^1 g \quad \Delta^2 g] = [g \quad q \quad G] \quad (4.45.a)$$

avec :

$$q(x) = \left( \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g \right) \quad (4.45.b)$$

et :

$$G = \left( \frac{\partial q}{\partial x} f - \frac{\partial f}{\partial x} q \right) + \left( \frac{\partial q}{\partial x} g - \frac{\partial g}{\partial x} q \right) u \quad (4.45.c)$$

Dans les deux régimes de fonctionnement du bioprocédé, les fonctions  $f$ ,  $g$ ,  $q$  et  $G$  possèdent les formes suivantes :

$$f = \begin{bmatrix} f_1 \\ f_2 \\ 0 \end{bmatrix} = \begin{bmatrix} r_X XW \\ r_A XW \\ 0 \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (4.46)$$

$$q = \begin{bmatrix} q_1 \\ q_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial W} \\ \frac{\partial f_2}{\partial W} \\ 0 \end{bmatrix}, \quad G = \begin{bmatrix} G_1 \\ G_2 \\ 0 \end{bmatrix} = \begin{bmatrix} G_{1,1} + G_{1,2}u \\ G_{2,1} + G_{2,2}u \\ 0 \end{bmatrix} \quad (4.47)$$

La matrice  $M$  s'écrit alors sous la forme :

$$M = \begin{bmatrix} 0 & q_1(x) & G_{1,1}(x) + G_{1,2}(x)u \\ 0 & q_2(x) & G_{2,1}(x) + G_{2,2}(x)u \\ 1 & 0 & 0 \end{bmatrix} \quad (4.48)$$

Les expressions des termes de cette matrice sont données en Annexe 2 pour chaque régime et/ou sous-régime de fonctionnement.

Le déterminant de cette matrice est de la forme :

$$\det(M) = \det \begin{bmatrix} q_1 & G_{1,1} + G_{1,2}u \\ q_2 & G_{2,1} + G_{2,2}u \end{bmatrix} = \underbrace{q_1 G_{2,1} - q_2 G_{1,1}}_{C_1(x)} + \underbrace{(q_1 G_{2,2} - q_2 G_{1,2})}_{C_2(x)} u \quad (4.49)$$

où  $C_1(x), C_2(x)$  sont des fonctions de  $G$  et de  $q$  et ne dépendent pas de  $u$ .

Le rang de la matrice  $M$  est inférieur à  $n$  si son déterminant (4.49) est nul.

Il faut alors distinguer deux cas (selon les développements du paragraphe 4.7.2.2) :

- si  $C_2(x) = 0$  : le rang de la matrice  $M$  dépend des conditions initiales mais pas de la commande  $u$ . Dans ce cas  $\sigma = \infty$ . Donc, il n'existe aucune région singulière,

- si  $C_2(x) \neq 0$  : l'entrée  $u$  apparaît explicitement dans  $G$  et par conséquent  $\sigma = 2$ , alors que le rang de  $M$  est de 2. Puisque  $\text{rang}(M) < \sigma + 1 < \infty$ , la singularité correspond à une surface et la commande optimale se calcule par  $u = -C_1(x)/C_2(x)$ .

Il est donc important d'étudier la fonction  $C_2(x)$  et plus précisément de savoir si elle peut s'annuler. Les expressions de cette fonction dépendent du régime de fonctionnement comme suit :

a. Régime oxydatif ( $\mu_2 = 0$ ) :

i. Sous-régime 1 :

$$C_2 = \frac{k_3}{k_2 k_1} A (S - S_{in}) \left( A \frac{\partial^2 q_{S,crit}}{\partial A^2} \frac{\partial q_S}{\partial S} - (S - S_{in}) \frac{\partial q_{S,crit}}{\partial A} \frac{\partial^2 q_S}{\partial S^2} \right) \quad (4.50)$$

ii. Sous-régime 2 :

$$C_2 = \frac{1}{k_1} A (S - S_{in}) \left( A \frac{\partial^2 q_{AC}}{\partial A^2} \frac{\partial q_S}{\partial S} - (S - S_{in}) \frac{\partial q_{AC}}{\partial A} \frac{\partial^2 q_S}{\partial S^2} \right) \quad (4.51)$$

b. Régime oxydo-fermentatif ( $\mu_3 = 0$ ) :

$$C_2 = \frac{k_3}{k_2 k_1} A (S - S_{in}) \left( A \frac{\partial^2 q_{S,crit}}{\partial A^2} \frac{\partial q_S}{\partial S} - (S - S_{in}) \frac{\partial q_{S,crit}}{\partial A} \frac{\partial^2 q_S}{\partial S^2} \right) \quad (4.52)$$

Dans les trois cas, la fonction  $C_2(x)$  dépend de la concentration en acétate et en substrat uniquement. On peut facilement constater que  $C_2(x) = 0$  pour  $A = 0$  ou  $S = S_{in}$ . A l'aide de simulations numériques, il a été vérifié que seules ces deux solutions existent.

La deuxième condition ( $S = S_{in}$ ) n'est pas retenue puisqu'elle induit des valeurs de substrat trop importante par rapport aux conditions opératoires d'une culture fed-batch des *E. coli*. Par ailleurs, pour  $A = 0$ , la solution est obligatoirement non singulière. Dans ce dernier cas, la solution optimale est soit aux limites, soit la commande maintenant le système sur la frontière des deux régimes (donnée par (4.44.d)).

Ainsi, l'expression de la commande singulière qui annule le déterminant de la matrice  $M$  est de la forme :

$$u_{\text{sing}} = \frac{q_2 G_{1,1} - q_1 G_{2,1}}{q_1 G_{2,2} - q_2 G_{1,2}} \quad (4.53)$$

#### 4.8.4 Récapitulatif

Finalement, la caractérisation de la solution optimale permet de déterminer les différentes expressions analytiques des solutions optimales possibles :

- $F_{in} = 0$  correspondant à un fonctionnement en batch,
- $F_{in} = F_{\text{max}}$ ,

- $F_{in} = u_{\text{sing}, R1}$  : arc singulier, valable pour un fonctionnement en régime oxydatif uniquement et incluant les deux sous-régimes,
- $F_{in} = u_{\text{sing}, R2}$  : arc singulier, valable pour un fonctionnement en régime oxydo-fermentatif uniquement,
- $F_{in} = u_{\text{path}}$  : solution non singulière correspondant à la commande maintenant la trajectoire du système sur la frontière  $q_S = q_{\text{Scrit}}$ .

Cependant, cette approche ne permet pas de déterminer la séquence optimale avec les instants de commutation. Pour ce faire, plusieurs scénarii doivent être considérés et pour chacun, les instants de commutation sont optimisés afin de maximiser la quantité de biomasse. Il s'agit donc d'un nouveau problème d'optimisation. La présence des arcs singuliers en fonction du régime de fonctionnement rend ce problème plus complexe comparé à d'autres cas de bioprocédés. En effet, dans le cas de la culture de levure [Renard, 2005] par exemple, seules trois solutions sont possibles : les valeurs limites inférieure, supérieure et la commande qui maintient la trajectoire sur la frontière (il n'existe pas de commande singulière).

## 4.9 Résultats numériques

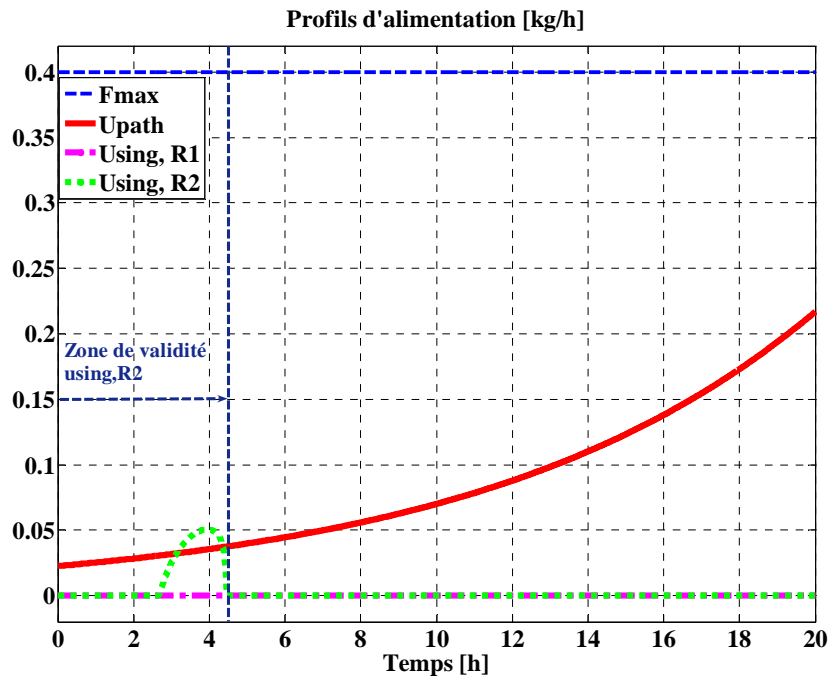
Le paragraphe précédent a permis de déterminer analytiquement les différentes lois de commandes optimales possibles. La solution optimale réelle est une combinaison de ces arcs optimaux. Il reste donc à optimiser le scénario d'alimentation optimal.

En pratique, une phase fed-batch est toujours précédée d'une phase batch. Celle-ci permet d'obtenir les conditions optimales du démarrage de la culture fed-batch. Ainsi, parmi les solutions présentées précédemment, seules celles qui sont non nulles sont prises en compte par la suite.

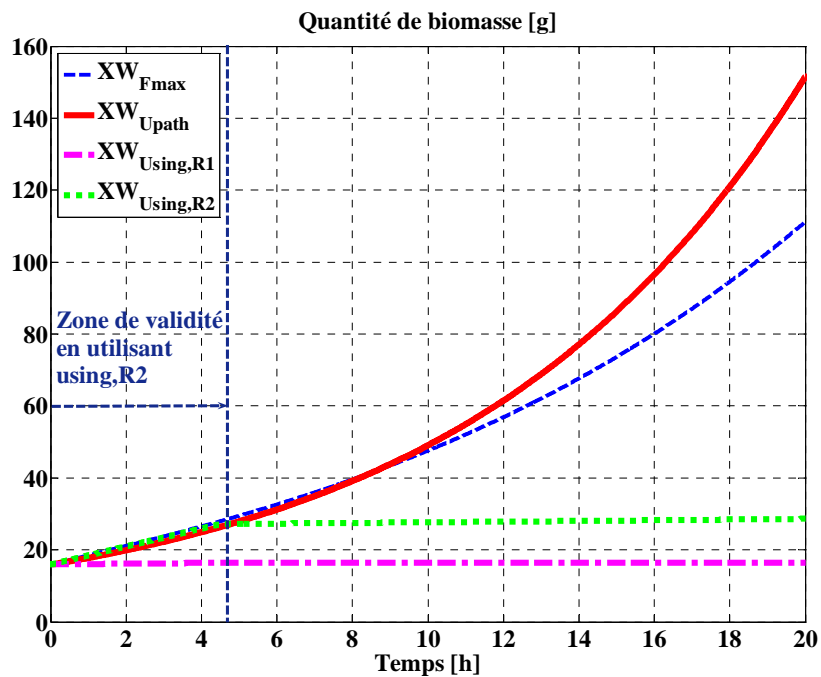
L'objectif de cette partie est de distinguer la solution optimale parmi les quatre solutions proposées en fed-batch. Pour cela, la Figure 4.2 présente l'évolution de la quantité  $X(t)W(t)$  pour les quatre profils d'alimentation ( $F_{\text{max}}, u_{\text{sing}, R1}, u_{\text{sing}, R2}, u_{\text{path}}$ ) présentés Figure 4.1. Pour chaque cas, les conditions initiales sont déterminées de façon à placer le système dans le régime adéquat (donc, les valeurs du substrat et d'acétate varient d'un cas à l'autre). Les valeurs initiales de la biomasse et du poids sont communes ( $X_0W_0 \approx 16g$ ). De plus, une étude des profils  $u_{\text{path}}$  maintenant le système sur la frontière permet d'établir que pour maximiser la croissance de la biomasse, l'optimum consiste à maintenir le système à la frontière pour  $A = 0$ . En effet, d'après (4.28), maximiser  $X(t)W(t)$  requiert d'avoir un taux spécifique de croissance de la biomasse  $r_X$  le plus élevé possible. Or, à la frontière, d'après (4.35)  $r_X = q_{S \text{crit}}/k_1$  et d'après (4.43.b),  $q_{S \text{crit}}$  est maximal pour  $A = 0$ . Aussi, sur les Figures 4.1 et 4.2, seule la commande  $u_{\text{path}}$  pour  $A = 0$  est prise en compte. La Figure 4.1 montre que :

- la commande non singulière  $u_{\text{path}}$  est une fonction exponentielle (comme établi par l'équation (4.44.d)), résultat également répandu chez les biologistes [Korz *et al.*, 1995],

- la commande singulière  $u_{\text{sing}, R1}$  correspond à une commande en batch dans ce cas. Ceci est dû aux contraintes sur les valeurs limites, à savoir que la commande doit être comprise entre 0 et  $F_{\text{max}}$ ,
- la commande singulière  $u_{\text{sing}, R2}$  a une zone de validité limitée. En effet, cette solution est calculée pour un fonctionnement en régime oxydo-fermentatif. Elle n'est donc valable que quand ce dernier est activé,
- La commande  $F_{\text{max}}$  alimente le bioprocédé avec le débit maximal de la pompe.

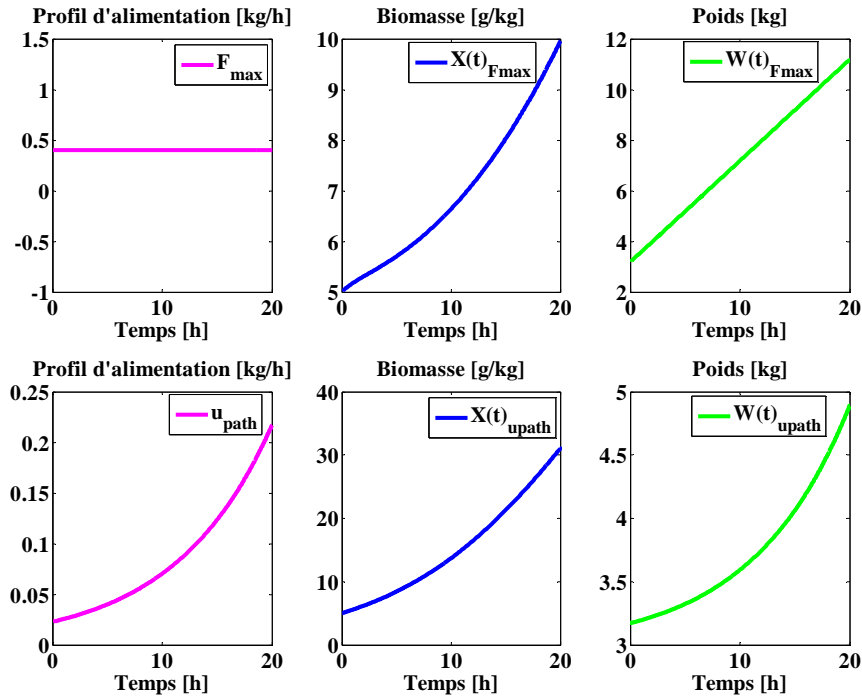


**Figure 4.1.** Tracés des quatre profils d'alimentation en fonction du temps.



**Figure 4.2.** Quantités de biomasse pour les quatre profils d'alimentation en fonction du temps.

La Figure 4.2 montre que la valeur de la quantité  $X(t_f)W(t_f)$  est maximale pour la commande non singulière  $u_{path}$ , elle est d'environ 150 g. Pour la commande  $F_{in} = F_{max}$ , la quantité de biomasse obtenue est égale à 111 g. En pratique, ce scénario n'est pas toujours réalisable. En effet, la Figure 4.3 montre qu'avec  $F_{in} = u_{path}$ ,  $W(t_f) = 4,83$  kg alors qu'avec  $F_{in} = F_{max}$ , le poids atteint environ 11 kg au bout de 20 h. Ce qui peut dépasser le poids maximal du bioprocédé.



**Figure 4.3.** Evolutions des concentrations en biomasse et des poids pour les deux profils d'alimentation :  $F_{max}$  et  $u_{path}$ .

Ainsi, pour le critère choisi, la commande  $u_{path}$  est la solution optimale, car elle permet de maximiser la quantité de biomasse à l'instant final ( $t_f = 20$  h). Ceci se confirme par nos connaissances biologiques concernant ce bioprocédé.

En effet, d'après la théorie du « goulot d'étranglement » (Figure 4.4), les conditions optimales du fonctionnement du bioprocédé sont définies quand la quantité de glucose correspond à la capacité oxydative. Dans ce cas, la concentration en acétate est nulle. Ce résultat a été établi, du point de vue biologique dans [Nakano *et al.*, 1997] et également établi dans le cas de la maximisation de la production de protéines dans [Ko *et al.*, 1994].

De ce fait, du point de vue biologique, le profil d'alimentation temporel  $F_{in}(t)$  maximise la quantité de biomasse si les conditions suivantes sont vérifiées :

- le système doit être placé à la frontière des deux régimes de fonctionnement ( $q_S = q_{Scrit}$ ), (Figure 4.5),
- la concentration en acétate doit être nulle ( $A = 0$ ).

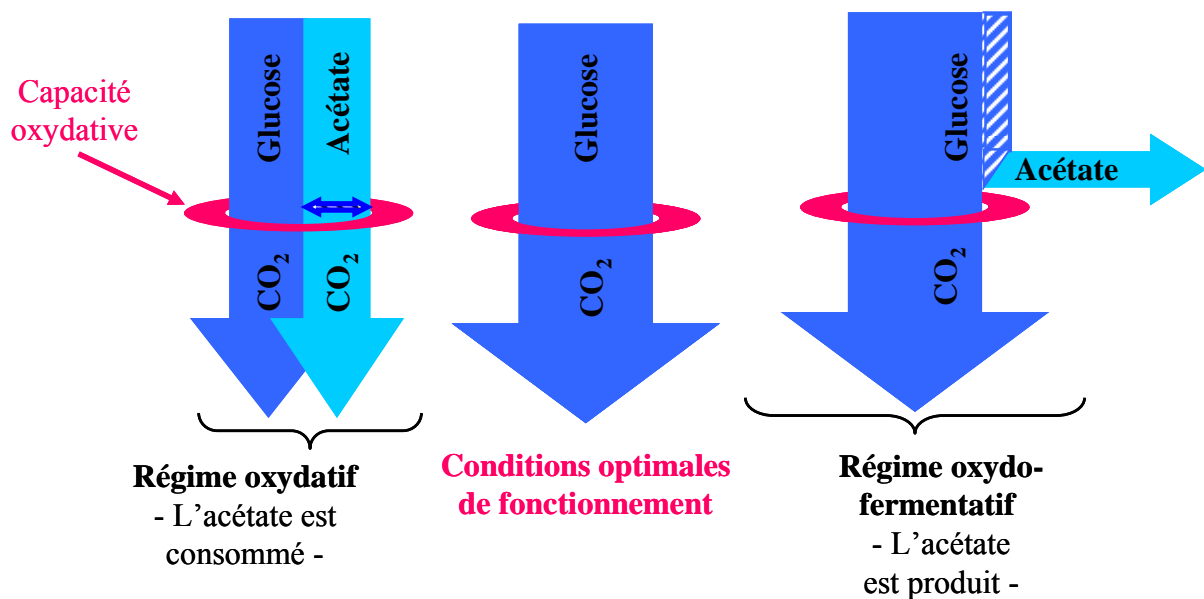


Figure 4.4. Représentation de la théorie du « goulot d'étranglement ».

### Frontière entre les deux régimes :

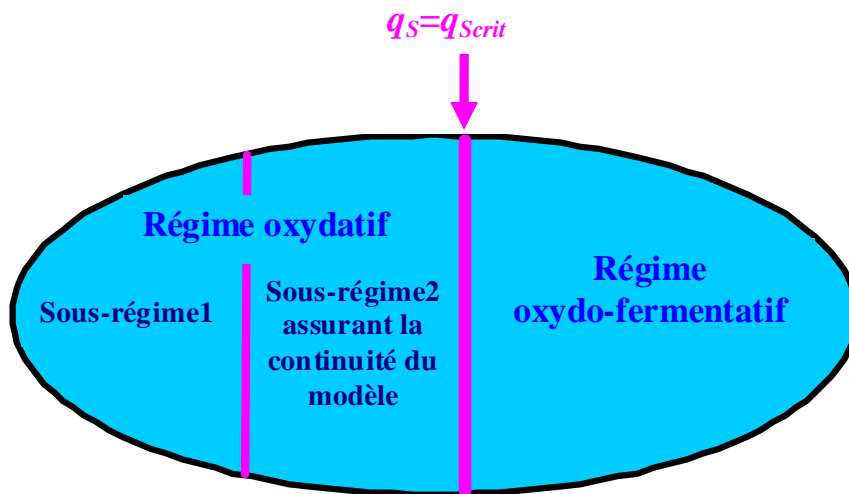


Figure 4.5. Schéma représentatif de la frontière entre les deux régimes de fonctionnement.

Finalement, la caractérisation de la solution optimale nous a permis de retrouver à l'aide d'outils mathématiques un résultat connu des biologistes. L'avantage des développements effectués dans ce chapitre réside dans la caractérisation analytique de la commande optimale pour ce type de procédé, ce qui permet de simplifier la phase de régulation autour de la trajectoire optimale.

## 4.10 Conclusions

L'objectif de ce chapitre était de déterminer un profil d'alimentation temporel optimal du bioprocédé de culture des bactéries *E. coli* afin de l'utiliser comme une trajectoire de

référence pour la commande prédictive non-linéaire développée au chapitre 5. Il s'agissait donc de résoudre un problème d'optimisation.

Pour ce faire, plusieurs méthodes de résolution numérique ont été présentées, induisant toutes cependant plusieurs inconvénients dans le cas des bioprocédés. Pour contrer ces inconvénients, la démarche développée par Srinivasan a été alors considérée et détaillée pour la classe des bioprocédés, puis appliquée dans le cas du bioprocédé de culture des *E. coli* examiné au chapitre 2. Après avoir déterminé les expressions analytiques des solutions optimales possibles et en choisissant de maximiser la quantité de biomasse à l'instant final, des résultats numériques ont été présentés pour caractériser la solution optimale en mode fed-batch. Ces résultats sont corroborés par une analyse biologique du système.

Les développements effectués dans ce chapitre ont permis avant tout de caractériser de façon analytique les différentes structures possibles pour la commande optimale du bioprocédé de culture des *E. coli*. Une étude en simulation a permis de sélectionner la trajectoire située sur la frontière des deux régimes comme étant celle permettant de maximiser la croissance de la biomasse. A ce stade, ce résultat n'a pas été recoupé par une procédure d'optimisation faisant intervenir toutes les structures possibles. Une perspective serait de mener à bien cette procédure en considérant explicitement dans la formulation la nature hybride du système.

Ainsi, il en ressort que le fonctionnement optimal d'une culture des *E. coli* consiste à réduire au maximum la concentration en acétate. De nombreuses études biologiques traitent de ce sujet dans la littérature. Globalement, trois familles d'approches sont utilisées :

- soit par la mise au point de procédé de culture, par exemple par des techniques de dialyses [Nakano *et al.*, 1997],
- soit par des manipulations génétiques des cellules *E. coli* [De Mey *et al.*, 2007],
- soit encore par des techniques de commande en mode fed-batch, principalement en régulant la concentration en substrat [Korz *et al.*, 1995]. Dans ce cas, la concentration de consigne en substrat est prise faible, à une valeur arbitraire, d'où le risque de sous-optimalité de la solution ainsi déterminée.

Dans [Eitman *et al.*, 2006], un état de l'art des différentes approches utilisées pour réduire la production de l'acétate est donné.

Ce travail de thèse s'associe à la troisième approche citée ci-dessus, puisque le but est de réguler la concentration de l'acétate à une valeur très faible, pour limiter le plus possible son effet inhibiteur. Dans ce sens, le profil optimal ainsi déterminé sera utilisé au chapitre 5 lors de la mise en œuvre de la stratégie de régulation de la concentration de l'acétate.



# 5. Mise en œuvre d'une stratégie de commande prédictive non-linéaire

## 5.1 Introduction

La commande des bioprocédés se limite très souvent dans l'industrie à la régulation de la température, du pH, de la concentration en oxygène dissous et en dioxyde de carbone dissous. La prise en compte d'autres constituants au sein de structures de commande est en effet encore très peu d'actualité. Or, la régulation des variables d'état est susceptible d'augmenter considérablement les performances du bioprocédé, à savoir son rendement et/ou sa productivité. Sachant que différentes cultures peuvent être développées dans un bioprocédé dans des domaines très variés – pharmaceutique, agroalimentaire, médical, biologique, traitement des eaux – la commande des bioprocédés constitue dès lors un véritable enjeu économique et en améliorer les performances un défi important de la communauté automatique.

De façon très générale, plusieurs techniques de commande avancées ont ainsi vu le jour dans ce domaine. Du point de vue méthodologique, Bastin et Dochain [Bastin et Dochain, 1990] ont proposé la mise en œuvre de la commande adaptative linéarisante pour les procédés biotechnologiques. La commande optimale a été ensuite implantée dans [Banga *et al.*, 1997]. En 2001, [Smets *et al.*, 2001] développent une commande optimale adaptative en combinant le meilleur des approches existantes (commande adaptative et commande optimale) afin d'améliorer la robustesse. Du point de vue applicatif, de nombreuses applications existent dans la littérature, on citera ici quelques références sans vouloir être exhaustif. Ainsi, dans [Queinnec *et al.*, 1992], la commande prédictive adaptative développée par M'Saad dans [M'Saad, 1987] a été appliquée à un bioprocédé de production d'alcool fonctionnant en mode fed-batch. La stratégie de commande développée dans [Bastin et Dochain, 1990] est appliquée dans [Rocha, 2003] pour la commande du bioprocédé de culture des *E. coli*. Depuis la fin des années 90, les automaticiens se sont intéressés à l'application de la commande prédictive aux bioprocédés. Dans [Rodrigues et Filho, 1999], elle a été appliquée à un bioprocédé de culture de pénicilline. [Renard, 2005] l'a implantée sous une forme RST robustifiée pour la commande d'un bioprocédé de culture de levures, montrant que cette stratégie pouvait se généraliser à d'autres types de bioprocédés.

Si l'on se restreint plus spécifiquement au type de cultures étudié dans cette thèse, on trouve également plusieurs travaux liés à la commande du bioprocédé de culture des bactéries *E. coli*. En effet, comme nous l'avons indiqué au chapitre 4, la commande de ce bioprocédé s'avère très intéressante dans le domaine pharmaceutique, car cette bactérie est très utilisée lors de la fabrication de médicaments et de vaccins.

Dans la littérature, la commande des bioprocédés de culture des bactéries *E. coli* est développée avec deux objectifs principaux. Le premier est primordial pour la vie bactérienne. Il s'agit d'assurer les conditions aérobiques dans le bioprocédé. Cet objectif est atteint en régulant la concentration en oxygène, par exemple à l'aide d'un correcteur par séquençement de gain en agissant sur le débit d'alimentation en substrat dans [Akesson *et al.*, 1999] et à l'aide d'une commande prédictive adaptative en agissant sur le taux d'agitation dans [Diaz *et*

*al.*, 1995]. Cette commande a été également utilisée pour atteindre ce même objectif dans [Diaz *et al.*, 1996] mais cette fois-ci en régulant simultanément la pression partielle en oxygène et en dioxyde de carbone et en agissant respectivement sur le taux d'agitation et le débit de circulation d'air. Le deuxième objectif est de maximiser la quantité de biomasse (bactéries *E. coli*). Dans ce sens, un régulateur issu de la méthodologie par modèle générique GMC (Generic Model Control) est développé dans [Jenzsch *et al.*, 2006], permettant de réguler le taux spécifique de croissance en biomasse en agissant sur le débit d'alimentation en substrat. Pour atteindre le même objectif, la stratégie de commande développée dans [Bastin et Dochain, 1990] est appliquée par Rocha [Rocha, 2003]. Il s'agit d'une loi de commande adaptative linéarisante, régulant la concentration en acétate en considérant le taux de dilution comme variable de commande.

Dans cette direction, la démarche proposée dans le cadre de ce travail et développée ci-dessous suppose que le premier objectif – régulation de la concentration en oxygène – est atteint en régulant la pression partielle en oxygène dissous à l'aide d'un régulateur PID et on s'intéresse au deuxième objectif qui consiste à maximiser la croissance de la biomasse. En effet, comme nous l'avons détaillé au chapitre 2, la concentration en acétate a un effet inhibiteur sur la croissance des bactéries. Il est primordial de réguler cette variable. De plus, nous avons montré au chapitre 4 que la quantité en biomasse se trouve maximisée si la concentration en acétate est nulle et si le profil d'alimentation appliqué correspond à une trajectoire de commande pré-spécifiée permettant au bioprocédé de se maintenir à la frontière entre les deux régimes de fonctionnement. Du point de vue de la stratégie de commande à élaborer, ceci se traduit par une régulation de la concentration en acétate à une valeur désirée (idéalement nulle) tout en garantissant le suivi d'une trajectoire de commande spécifique.

Or, comme on a pu le percevoir lors des chapitres précédents, le bioprocédé *E. coli* est un système complexe, fortement non-linéaire et contraint. Ceci justifie le recours à une technique de commande avancée prenant en compte la dynamique du système et ses différentes contraintes. Dans ce sens, des techniques prédictives non-linéaires, relativement simples à implanter et robustes, connaissent un succès certain dans le domaine des procédés chimiques. Ainsi, dans [Al Seyab et Cao, 2006], la commande prédictive non-linéaire est appliquée à un générateur de gaz de la société ALSTOM. Elle est utilisée pour la régulation de la température d'un réacteur chimique industriel dans [Xaumier *et al.*, 2002] et pour la commande des colonnes à distiller dans [Balasubramhanya et Doyle III, 2000] et [Kawathekar et Riggs, 2007]. Grâce aux résultats très intéressants obtenus lors de l'application de la commande prédictive non-linéaire aux procédés chimiques, quelques applications aux bioprocédés commencent à apparaître. On citera par exemple son utilisation pour la régulation de la température d'un bioprocédé de culture de levures [Nagy, 2007], les auteurs montrant de plus l'avantage d'appliquer une commande prédictive non-linéaire comparativement à une commande prédictive linéaire.

Les travaux de thèse se sont donc orientés vers l'élaboration d'une stratégie de commande prédictive non-linéaire adaptée au bioprocédé *E. coli*, afin de réguler la concentration en acétate tout en forçant le débit d'alimentation à suivre une trajectoire de référence. Les performances obtenues en simulation seront comparées à celles fournies par des stratégies déjà étudiées dans la littérature.

Pour cela, une première partie présente le principe et l'historique de la commande prédictive ainsi que le cadre théorique dans lequel s'insère l'algorithme de la commande prédictive non-linéaire. Préalable nécessaire à la mise en œuvre d'une structure de commande sous

formalisme d'état, une deuxième partie se penche sur l'étude de la commandabilité du système en se basant sur le modèle d'état défini lors des chapitres précédents. Une troisième partie détaille l'application de la stratégie prédictive non-linéaire au bioprocédé *E. coli*, rappelant l'objectif à atteindre pour la culture, présentant le critère considéré ainsi que la méthode de résolution envisagée à partir de techniques de CVP (Control Vector Parametrization). Les résultats obtenus dans ce cadre sont ensuite analysés avant d'être comparés dans une dernière partie à ceux que fournissent deux autres types de commande, la commande adaptative linéarisante et la commande générique par modèle.

## 5.2 Commande prédictive

La commande prédictive fait partie de l'ensemble des techniques de commandes avancées apparues récemment dans le domaine de l'Automatique afin de répondre à un réel besoin du monde industriel en termes de performances élevées des boucles d'asservissement des systèmes industriels complexes. Ces performances doivent être atteintes tout en respectant des contraintes de fonctionnement et de production toujours plus élevées.

Ce paragraphe a pour objectif de résumer les principes fédérant l'ensemble des techniques de commande prédictive, de proposer un bref historique de cette approche et enfin de définir le cadre mathématique dans lequel sera développé par la suite la structure de commande prédictive non-linéaire élaborée pour le bioprocédé *E. coli*.

### 5.2.1 Principe de la commande prédictive

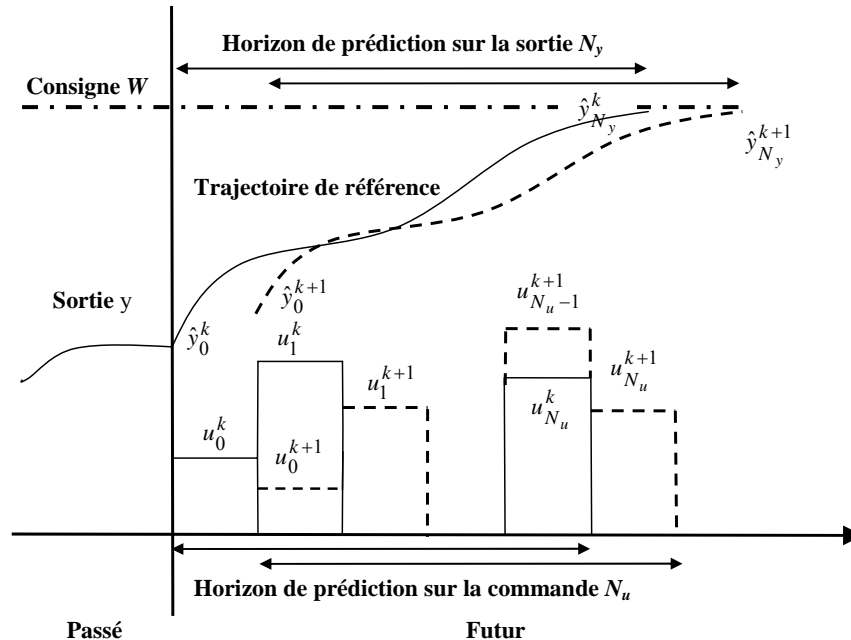
Le principe de la commande prédictive consiste à créer pour le système asservi un effet anticipatif par rapport à une trajectoire à suivre connue à l'avance, en se basant sur la prédiction du comportement futur du système et en minimisant l'écart de ces prédictions à la trajectoire au sens d'une certaine fonction coût, tout en respectant des contraintes de fonctionnement. Cette idée est simple et intuitive, pratiquée de façon assez systématique dans la vie quotidienne. Par exemple, un piéton analyse l'état de son chemin sur un horizon assez lointain pour déterminer le chemin le plus rapide à prendre, en tenant compte des différentes contraintes (« feux de circulation », « éviter la foule », « passage glissant » ...).

Les techniques de commande prédictive s'appuient sur un modèle de prédiction déterminé hors ligne. Cette particularité permet de classer la commande prédictive dans la grande famille des commandes à base de modèles, dite 'MBC' (Model Based Control). A partir d'une trajectoire de référence à suivre connue, elles réalisent en temps réel à chaque période d'échantillonnage les étapes suivantes :

1. Calculer les prédictions des variables de sortie  $\hat{y}$  sur un horizon de prédiction sur la sortie  $N_y$ ,
2. Minimiser un critère quadratique à horizon fini portant sur les erreurs de prédictions futures, écarts entre la sortie prédite du système et la consigne future,
3. Obtenir une séquence de commandes futures sur un horizon de commande  $N_u$  pouvant être différent de  $N_y$ ,

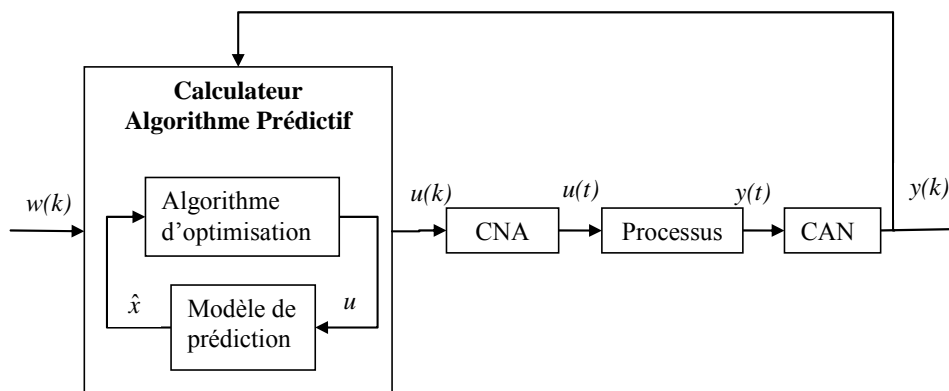
4. Appliquer uniquement la première valeur de cette séquence sur le système,
5. Répéter ces étapes à la période d'échantillonnage suivante, selon le principe de l'horizon fuyant.

L'ensemble est résumé sur la Figure 5.1.



**Figure 5.1.** Principe de l'horizon fuyant.

Le principe de la commande prédictive permet d'établir le schéma de fonctionnement présenté Figure 5.2.



**Figure 5.2.** Principe de fonctionnement d'un algorithme prédictif.

#### Remarque :

La détermination de la loi de commande prédictive se fait par résolution, à chaque pas d'échantillonnage, d'un problème de commande optimale à horizon fini.

Dans le cas général d'un système sous contraintes, la minimisation du critère prédictif requiert la résolution effective d'un problème d'optimisation en ligne. Seule la commande prédictive des systèmes linéaires invariants dans le temps, restreinte au cas sans contraintes, ne requiert

pas la résolution effective de ce problème d'optimisation en ligne, car le correcteur est à son tour linéaire invariant et sa description analytique peut être obtenue hors-ligne [Boucher et Dumur, 1996]. Un avantage primordial issu de cette caractéristique est que toute la théorie de la commande linéaire peut être utilisée pour l'analyse et la synthèse de la loi prédictive sans contraintes [Garcia et Morari, 1982]. Clarke [Clarke *et al.*, 1987a] a ainsi montré que dans ce cas le réglage des performances d'une boucle de commande prédictive s'effectue de façon claire par le choix des paramètres de réglage de la loi prédictive, horizons de prédiction ou pondérations.

Malheureusement, le procédé envisagé dans ce travail, non-linéaire et sous contraintes, implique qu'il faille plutôt s'orienter vers des techniques prédictives non-linéaires nécessitant la mise en œuvre d'une stratégie d'optimisation, comme envisagé ci-dessous.

## 5.2.2 Bref historique de la commande prédictive

Issues de la même philosophie, plusieurs variantes de la commande prédictive ont été développées dans la littérature depuis son apparition. Ce paragraphe dresse un bilan non exhaustif de cette évolution.

La philosophie de la commande prédictive est régulièrement citée depuis la fin des années 60 par les spécialistes de la commande optimale. Ainsi, la lecture de cette citation traduite de [Lee et Markus, 1967] montre que l'idée était effectivement présente depuis relativement longtemps :

« Une façon d'obtenir un bouclage à partir de la connaissance des solutions en boucle ouverte est de mesurer l'état courant du système et de calculer très rapidement la trajectoire optimale. La première portion de celle-ci est alors appliquée pendant une courte période à la fin de laquelle une nouvelle mesure est acquise et une nouvelle solution optimale est calculée. La procédure est ainsi répétée ».

Mais ce n'est qu'au milieu des années 70 qu'elle a été mise en œuvre dans l'industrie grâce notamment à Richalet *et al.* qui ont développé le logiciel IDCOM (Identification-Commande) [Richalet *et al.*, 1976]. Les premières applications industrielles de la commande prédictive ont été réalisées dans le domaine pétrolier et pétrochimique [Richalet *et al.*, 1978].

En 1979 et dans le même domaine, des ingénieurs de Shell, Cutler et Ramaker, ont présenté leur expérience sur un craqueur catalytique<sup>7</sup> [Cutler et Ramaker, 1979], en se basant sur l'approche dite DMC (Dynamic Matrix Control).

En 1987, Clarke et son équipe d'Oxford ont proposé la première version de la commande prédictive généralisée GPC (Generalized Predictive Control) [Clarke *et al.*, 1987a] et [Clarke *et al.*, 1987b]. La même année, Richalet *et al.* ont développé une autre variante de la commande prédictive, dénommée commande prédictive fonctionnelle [Richalet *et al.*, 1987], en l'appliquant notamment au domaine de la robotique.

A partir des années 90, le nombre d'applications de la commande prédictive a explosé (tout d'abord aux Etats-Unis, puis au Japon et ensuite en Europe). De nombreuses applications sont présentées dans la littérature, principalement dans le domaine des procédés chimiques [Bequette, 1991], [De Keyser, 1998], [Camacho et Bordons, 1998] et [Morari et Lee, 1999], dans le domaine de la robotique [Company et Pierrot, 1999], [Essen et Nijmeijer, 2001] et

<sup>7</sup> Unité de raffinerie qui permet de transformer par l'action d'un catalyseur les fractions lourdes du pétrole brut en fractions plus légères nécessaires à la fabrication des essences et gazole.

dans le domaine des bioprocédés [Rodrigues et Filho, 1999], [Zhu *et al.*, 2000] et [Ramaswamy *et al.*, 2005].

### 5.2.3 Commande prédictive non-linéaire

Le problème simple issu de la minimisation du critère prédictif dans le cas d'un système linéaire invariant sans contrainte se complique singulièrement lorsque l'on envisage le cas de la commande prédictive d'un système non-linéaire sous contraintes. En ce qui concerne les contraintes en particulier, la meilleure solution consiste alors à considérer les contraintes depuis l'étape de synthèse, imposant ainsi leur présence au sein du problème d'optimisation. Cette façon de voir la commande prédictive sous contraintes engendre notamment la difficulté de mise en forme appropriée des contraintes induisant une forme canonique utilisable par les programmes d'optimisation. Cette problématique est très largement développée par exemple dans [Camacho et Bordons, 2003], [Maciejowski, 2002] et [Rossiter, 2003], références dans lesquelles on trouve également des réponses aux questions de mise en œuvre. La stratégie d'optimisation engendre également l'incertitude quant à la stabilité de la structure asservie résultante. Celle-ci étant globalement non-linéaire, tous les outils classiques d'analyse de stabilité des systèmes linéaires sont inapplicables. Des outils spécifiques doivent être utilisés, basés essentiellement sur la méthode de Lyapunov [Keerthi et Gilbert, 1988], [Mayne et Michalska, 1990]. A ce sujet, [Mayne *et al.*, 2000] propose une vision synthétique complète des conditions de stabilité de schémas de commande prédictive non-linéaire.

De façon générale, la commande prédictive non-linéaire est appliquée aux systèmes non-linéaires en se basant pour la prédiction notamment sur un modèle non-linéaire du système, évitant ainsi le recours à une linéarisation préalable. Comme on l'a vu, elle repose sur la résolution d'un problème d'optimisation à chaque pas d'échantillonnage. Ce dernier consiste à minimiser une fonction coût sur un horizon de prédiction. Une formulation très globale de cet ensemble est donnée ci-dessous.

#### 5.2.3.1 Définition de la fonction coût

Considérons un système non-linéaire discret décrit par les équations aux différences suivantes :

$$\begin{aligned} x_{k+1} &= \mathcal{F}_d(x_k, u_k) \\ y_k &= \mathcal{G}_d(x_k, u_k) \end{aligned} \quad (5.1)$$

où :

- $x$  est l'état du système,
- $k$  est l'indice associé à une période d'échantillonnage supposée constante,
- $u$  est l'entrée de commande,
- $\mathcal{F}_d$  et  $\mathcal{G}_d$  sont des fonctions issues du modèle dynamique du bioprocédé,
- $y$  est la sortie à réguler.

Lorsque le système se trouve à l'état  $x$  et à l'instant d'échantillonnage  $k$ , une fonction coût est associée à chaque profil de commande  $u$ . La fonction coût intervenant dans l'algorithme de la commande prédictive non-linéaire est une fonction de la forme :

$$J = \sum_{i=0}^{N-1} L(x_{k+i}, u_{k+i}) + F(x_N) \quad (5.2a)$$

avec  $L(x_k, u)$  le coût instantané pour l'étape prédite,  $N = N_y$  le nombre d'étapes prises en compte pour la prédiction et  $F(x_N)$  un coût terminal que l'on peut éventuellement imposer. Le coût instantané est souvent choisi quadratique ; ainsi, l'expression (5.2.a) peut s'exprimer par :

$$J = x_N^T P_w x_N + \sum_{i=0}^{N-1} ((x_{set_{k+i}} - x_{k+i})^T Q_w (x_{set_{k+i}} - x_{k+i}) + u_{k+i}^T R_w u_{k+i}) \quad (5.2b)$$

$P_w$  et  $Q_w$  sont des matrices semi-définies positives,  $R_w$  est une matrice définie positive. Toutes ces matrices constituent des pondérations soit pour le coût instantané, soit pour le coût terminal.

### 5.2.3.2 Formulation du problème d'optimisation

Compte-tenu du critère précédent, l'algorithme permettant l'élaboration de la loi de commande prédictive non-linéaire se trouve formalisé par les étapes suivantes.

1. Calculer la séquence optimale  $\chi = \{x_{k+1}, \dots, x_{k+N}, u_k, \dots, u_{k+N-1}\}$  à l'instant  $k$  qui minimise la fonction quadratique  $J$  :

$$\chi_{opt} = \arg \min_{\chi} J \quad (5.3.a)$$

sous les contraintes suivantes ( $k = \overline{0, N-1}$ ) :

- contraintes sur le modèle : 
$$\begin{cases} x_{k+1} = \mathcal{F}_d(x_k, u_k) \\ (x_0 \text{ connu}) \end{cases} \quad (5.3.b)$$

- contraintes sur les états : 
$$\underline{p} \leq P \hat{x}_k \leq \bar{p} \quad (5.3.c)$$

- contraintes sur la commande : 
$$\underline{r} \leq R u_k \leq \bar{r} \quad (5.3.d)$$

- contraintes terminales éventuelles : 
$$p_T \leq P_T x_N \quad (5.3.e)$$

$P$ ,  $P_T$  et  $R$  sont des matrices constantes.  $\underline{p}$ ,  $\bar{p}$ ,  $\underline{r}$ ,  $\bar{r}$  sont des vecteurs constants vérifiant les inégalités  $\underline{p} \leq \bar{p}$  et  $\underline{r} \leq \bar{r}$ .

2. Obtenir la séquence de commande optimale  $\{u_{k/k}, \dots, u_{k+N-1/k}\}$
3. Appliquer la première commande  $u_k = u_{k/k}$  au système
4. Répéter la procédure en posant  $k = k + 1$  selon le principe de l'horizon fuyant

Le problème ainsi formulé de façon très générale est appliqué et adapté dans les paragraphes suivants au cas du bioprocédé *E. coli*.

## 5.3 Etude de la commandabilité du bioprocédé *E. coli*

Dès lors que l'on souhaite mettre en œuvre une commande par retour d'état ou faisant intervenir l'état du système, il est nécessaire avant d'effectuer la synthèse de la loi de commande de s'assurer que le système est commandable. Pour cela, ce paragraphe rappelle dans un premier temps les conditions à vérifier garantissant la commandabilité d'un système non-linéaire affine en la commande (ce qui correspond au modèle du bioprocédé *E. coli*) puis développe ces conditions pour le système plus particulièrement étudié.

### 5.3.1 Conditions de commandabilité

Un système est dit complètement commandable si, par action sur l'entrée de commande, n'importe quel point de l'espace d'état est atteint en un temps fini.

Soit un système non-linéaire s'écrivant sous la forme suivante :

$$\dot{x} = f(x) + g(x)u \quad (5.4.a)$$

avec les champs de vecteurs  $f$  et  $g$  suivants :

$$f(x) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix} \quad \text{et} \quad g(x) = \begin{pmatrix} g_1(x_1, \dots, x_n) \\ \vdots \\ g_n(x_1, \dots, x_n) \end{pmatrix} \quad (5.4.b)$$

où  $f$  et  $g$  sont des fonctions non-linéaires en les états.

Une condition nécessaire et suffisante pour que le système (5.4) soit complètement commandable est que l'algèbre de Lie du système, donc une sous-algèbre de Lie engendrée par la famille  $E$  des champs de vecteurs  $f$  et  $g$ , soit de dimension égale à  $n$  [Sussmann et Jurdjevic, 1972]. Ceci revient à écrire :

$$\dim L(E) = n \quad \text{avec} \quad L(E) = \{g \quad [f, g] \quad [f, [f, g]] \quad [f, [f, [f, g]]] \quad \dots\}$$

$$\text{avec} \quad [f, g] = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g.$$

La matrice de commandabilité [Sussmann et Jurdjevic, 1972] est alors définie comme suit :

$$\mathcal{C} = [g \quad [f, g] \quad [f, [f, g]] \quad \dots].$$

Le système (5.4) est commandable si et seulement si le rang de la matrice  $\mathcal{C}$  est égal à sa dimension  $n$ .

### 5.3.2 Commandabilité du bioprocédé *E. coli*

L'objectif est d'appliquer les conditions de commandabilité présentées ci-dessus au modèle d'état non-linéaire du bioprocédé *E. coli*. On se base comme pour les chapitres précédents sur les valeurs numériques du bioréacteur présenté à la fin du chapitre 2.

Reprenons alors le modèle dynamique du bioprocédé de culture des *E. coli* présenté au chapitre 2 :

$$\begin{aligned}
 \frac{dX}{dt} &= r_X X - \frac{F_{in}}{W} X \\
 \frac{dS}{dt} &= -q_S X + \frac{F_{in}}{W} (S_{in} - S) \\
 \frac{dA}{dt} &= r_A X - \frac{F_{in}}{W} A \\
 \frac{dW}{dt} &= F_{in}
 \end{aligned} \tag{5.5}$$

où :

- $X$ ,  $S$  et  $A$  sont respectivement les concentrations en biomasse, en substrat (glucose) et en acétate et  $W$  est le poids,
- $S_{in}$  représente la concentration en substrat dans le milieu d'alimentation,
- $F_{in}$  est le débit d'alimentation en substrat, représentant la commande  $u$ ,
- $r_X$ ,  $r_A$  et  $q_S$  sont respectivement les taux spécifiques de croissance, de consommation d'acétate et de consommation de substrat. Ils dépendent des variables d'état et sont décrits au chapitre 2. Les expressions de  $r_X$  et  $r_A$  dépendent des régimes et/ou sous-régimes de fonctionnement. Elles sont rappelées dans le Tableau 5.1.

Taux	Régime oxydatif		Régime oxydo-fermentatif
	Sous-régime 1	Sous-régime 2	
$r_X$	$\frac{q_S}{k_1} + \frac{q_{AC}}{k_4}$	$\frac{q_S}{k_1} + \frac{k_{os}}{k_4} \frac{q_{Scrit} - q_S}{k_{oa}}$	$\frac{q_S - q_{S,crit}}{k_2} + \frac{q_{S,crit}}{k_1}$
$r_A$	$-q_{AC}$	$k_{os} \frac{q_S - q_{Scrit}}{k_{oa}}$	$k_3 \frac{q_S - q_{S,crit}}{k_2}$

**Tableau 5.1.** Expressions des taux  $r_X$  et  $r_A$  selon le régime de fonctionnement.

Le système d'équations (5.5) peut s'écrire sous la forme (5.4) en adoptant les notations :

$$\begin{aligned}
 u &= F_{in} \\
 x &= [X \ S \ A \ W]^T \\
 f &= [r_X X \ -q_S X \ r_A X \ 0]^T \\
 g &= [-X/W \ (S_{in} - S)/W \ -A/W \ 1]^T
 \end{aligned}$$

Il s'agit bien d'un système non-linéaire et affine en la commande. Puisque le système étudié (5.5) est de dimension 4, la matrice de commandabilité s'écrit sous la forme :

$$\mathcal{C} = [g \ [f, g] \ [f, [f, g]] \ [f, [f, [f, g]]] ] \tag{5.6}$$

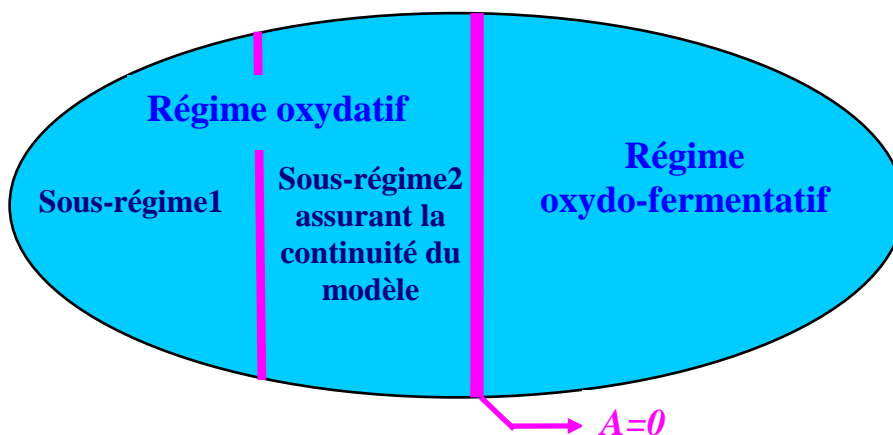
Le système est commandable si le rang de cette matrice est égal à la dimension du système, autrement dit si la matrice  $\mathcal{C}$  est inversible. Un calcul analytique complet s'avère très difficile à mettre en œuvre compte-tenu de la complexité des expressions, en particulier des taux spécifiques. Pour contourner cette difficulté, le déterminant de la matrice  $\mathcal{C}$  a été calculé à l'aide de la *Symbolic Math Toolbox* de *Matlab*<sup>TM</sup> pour les trois cas de fonctionnement du bioprocédé. Les résultats sont regroupés dans le Tableau 5.2.

Régime de fonctionnement		Déterminant de la matrice de commandabilité $\mathcal{C}$
Régime oxydatif	Sous-régime 1	$\frac{-1,25 \cdot 10^{23} X^2 A}{(357 + 2500 S)^4 (1309 + 2500 S)(117 + 20 A)}$
	Sous-régime 2	$\frac{-4 \cdot 10^{23} (-1,2 X + 1,3 + 0,19 A) X^2}{(357 + 2500 S)(869 + 125 A)^2}$
Régime oxydo-fermentatif		$\frac{3,5 \cdot 10^{22} (0,25 A + 1,7 - 2,7 X) X^2}{(357 + 2500 S)^4 (869 + 125 A)^2}$

**Tableau 5.2.** Déterminant de la matrice  $\mathcal{C}$  selon le régime de fonctionnement.

Déterminons alors les points pour lesquels il s'annule dans chaque cas.

- Pour le sous-régime 1 représentant une partie du régime oxydatif, il est clair que le déterminant de la matrice  $\mathcal{C}$  s'annule si au moins l'une des concentrations – soit en biomasse  $X$  soit en acétate  $A$  – est nulle. Or, la concentration en biomasse  $X$  est toujours non nulle car pour une culture de bactéries, il est essentiel de mettre une quantité de biomasse à l'état initial. Cette biomasse n'est autre qu'un ensemble de bactéries *E. coli*. Le déterminant pourrait alors s'annuler pour  $A = 0$ . Mais ce point n'est pas atteignable dans ce cas comme le montre la Figure 5.3. En effet, la concentration en acétate est nulle uniquement sur la frontière des deux régimes de fonctionnement du bioprocédé où l'acétate n'est ni consommé ni produit.



**Figure 5.3.** Schéma représentatif des différents cas de fonctionnement.

- Pour le sous-régime 2, le déterminant de la matrice de commandabilité s'annule pour les couples  $(X, A)$  tels que  $X = 1,08 + 0,15A$ . Or, du fait de l'effet d'inhibition provoqué par la présence d'acétate, on ne peut se trouver réellement qu'en présence de faibles valeurs de concentration en acétate. L'application de la formule précédente montre qu'il faudrait alors une faible valeur de concentration en biomasse pour que la relation s'annule, ce qui n'est pas réaliste ici. En effet, on introduit toujours une quantité de biomasse initiale non négligeable et l'objectif est ensuite de faire croître cette biomasse. Ainsi, l'ensemble des points pour lesquels le déterminant de la matrice de commandabilité est nul n'est pas réalisable.
- Pour le régime oxydo-fermentatif, le déterminant de la matrice  $\mathcal{C}$  s'annule pour les couples  $(X, A)$  tels que  $X = 0,63 + 0,09A$ . Le même raisonnement que précédemment montre que ces couples ne sont pas réalisables.

Cette analyse permet donc de déduire que le bioprocédé de culture des *E. coli* modélisé par la structure d'état non-linéaire affine en la commande (5.5) est totalement commandable dans le cadre des conditions opératoires usuelles de la culture des bactéries *E. coli*.

## 5.4 Application de la commande prédictive non-linéaire au bioprocédé *E. coli*

Le but de ce paragraphe est la mise en œuvre d'une structure de commande prédictive non-linéaire pour le bioprocédé *E. coli*, en se basant sur le formalisme théorique défini précédemment, avec malgré tout une adaptation à l'objectif de commande visé et à la structure du système. Après avoir rappelé cet objectif, on envisage successivement la formulation du critère et la résolution du problème d'optimisation.

### 5.4.1 Objectif de commande

Notre objectif est de maximiser la quantité de biomasse, et donc la croissance de la biomasse. Il a été montré au chapitre 4 que, pour atteindre cet objectif, deux conditions devaient être satisfaites :

- l'acétate ne doit être ni consommé ni produit. La concentration en acétate doit rester nulle  $A = 0$  g/kg ,
- le système doit être maintenu à la frontière des deux régimes (oxydatif et oxydo-fermentatif), ce qui se traduit en termes de taux spécifique par  $q_S = q_{Scrit}$ . Le profil de débit d'alimentation permettant de satisfaire cette condition a été déterminé au chapitre 4 sous la notation  $u_{path}$ .

Ainsi, pour maximiser la quantité de biomasse, la concentration en acétate doit être régulée à  $A_{set} = 0$  g/kg et le débit d'alimentation  $F_{in}$  doit suivre la trajectoire  $u_{path}|_{A=0}$ , relation (4.44.d).

Globalement donc, l'objectif visé par la stratégie de commande prédictive non-linéaire sera de réguler la concentration en acétate  $A$  à une valeur de référence  $A_{set}$ , tout en forçant le débit d'alimentation  $F_{in}$  à suivre un profil d'alimentation de référence noté par la suite  $F_{ref}$ .

## 5.4.2 Formulation de la commande prédictive non-linéaire appliquée au bioprocédé *E. coli*

Le point de départ de cette formulation est constitué des relations introduites dans un contexte général au paragraphe 5.2.3, qu'il est nécessaire d'adapter ici au bioprocédé de culture des *E. coli*.

La structure du modèle rappelée par (5.5) peut s'écrire sous la forme suivante :

$$\begin{aligned}\dot{x}(t) &= \mathcal{F}(x(t), F_{in}(t)) \\ A(t) &= H x(t)\end{aligned}\quad (5.7.a)$$

où :

- $x = [X \ S \ A \ W]^T$  est le vecteur des variables d'état,
- $\mathcal{F}$  est une fonction non-linéaire incluant toutes les relations présentes dans (5.5),
- $H = [0 \ 0 \ 1 \ 0]$  est la matrice d'observation de la représentation d'état liant la grandeur de sortie à réguler – la concentration en acétate  $A$  – au vecteur d'état.

ou plus précisément si l'on se réfère à la forme affine en la commande (5.4.a) :

$$\dot{x}(t) = \underbrace{\begin{bmatrix} r_X X \\ -q_S X \\ r_A X \\ 0 \end{bmatrix}}_{f(x)} + \underbrace{\begin{bmatrix} -X/W \\ (S_{in} - S)/W \\ -A/W \\ 1 \end{bmatrix}}_{g(x)} \underbrace{\underbrace{F_{in}(t)}_{u(t)}}_{u(t)} \quad (5.7.b)$$

On supposera dans toute la suite que l'ensemble du vecteur d'état est mesurable.

Par la suite, dans la perspective de la mise en œuvre d'une commande prédictive, il est a priori nécessaire de disposer d'un modèle discret du système pour la prédiction. On considèrera donc que ce modèle discret issu de (5.7.a) par discrétisation à une période d'échantillonnage  $T_e$  est donné par la relation :

$$\begin{aligned}x(k+1) &= \mathcal{F}_{dis}(x(k), F_{in}(k)) \\ A(k) &= H x(k)\end{aligned}\quad (5.7.c)$$

Le paragraphe 5.4.3 reviendra sur l'opportunité d'un tel modèle et envisagera une solution alternative.

### 5.4.2.1 Définition du critère

L'objectif étant de suivre une valeur de concentration d'acétate spécifiée ainsi qu'un profil de commande de référence (réguler la concentration en acétate  $A$  à une valeur de référence  $A_{set}$ , tout en forçant le débit d'alimentation  $F_{in}$  à suivre un profil de référence  $F_{ref}$ ), la structure générale des critères (5.2) est modifiée de façon à s'orienter vers une formulation similaire à l'algorithme de commande prédictive généralisée à modèles de référence multiples (GPC/MRM) introduit dans le cas linéaire [Irving *et al.*, 1986]. On s'oriente donc dans ce cas vers un régulateur introduisant un bouclage sur la sortie  $A$  plutôt que sur l'état tout entier, et

l'on n'introduira pas de coût terminal ni de contraintes terminales. Dans ce cas, la fonction coût quadratique  $J$  prend la forme suivante :

$$J = \sum_{j=1}^N (A_{set_{k+j}} - \hat{A}_{k+j})^2 + \eta \sum_{j=1}^N (F_{ref_{k+j-1}} - F_{in_{k+j-1}})^2 \quad (5.8)$$

où l'horizon de prédiction supérieur sur la sortie  $N_y$  est égal à l'horizon sur la commande  $N_u$  :  $N = N_y = N_u$ .  $\hat{A}$  correspond à la prédiction de la concentration en acétate effectuée par l'intermédiaire du modèle (5.7.c).  $\eta$  est un facteur de pondération sur la commande.

Compte-tenu de ce critère, le problème d'optimisation est alors formulé comme suit :

$$\min_{\chi} \sum_{j=1}^N (A_{set_{k+j}} - \hat{A}_{k+j})^2 + \eta \sum_{j=1}^N (F_{ref_{k+j-1}} - F_{in_{k+j-1}})^2$$

sous les contraintes :

$$\begin{cases} \hat{A}_{k+1} = H \mathcal{F}_{dis}(x_k, F_{in_k}) \\ \vdots \\ \hat{A}_{k+N} = H \mathcal{F}_{dis}(\hat{x}_{k+N-1}, F_{in_{k+N-1}}) \\ F_{in_k} \geq 0, \quad \forall k \in \mathcal{N} \\ x_k \geq 0, \quad \forall k \in \mathcal{N} \end{cases} \quad (5.9)$$

où  $\chi = \{F_{in_k}, \dots, F_{in_{k+N-1}}, \hat{A}_{k+1}, \dots, \hat{A}_{k+N}\}$  est le vecteur d'optimisation correspondant.

#### 5.4.2.2 Prise en compte d'une erreur de modèle

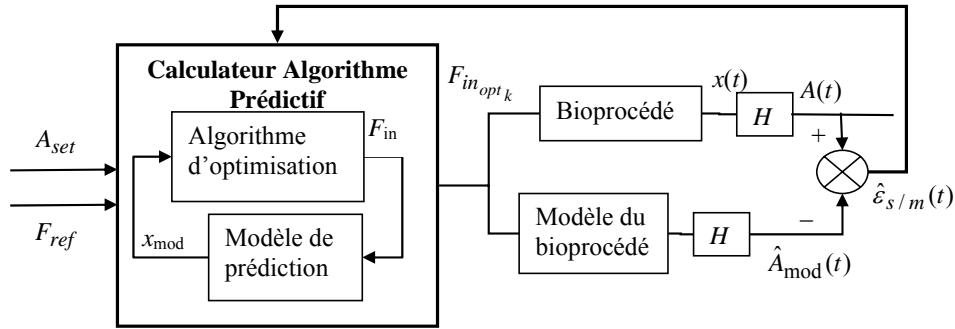
La loi de commande qui se déduit du problème d'optimisation précédent n'inclut pas d'action intégrale, et toute désadaptation entre le modèle et le système introduit une erreur au niveau des performances de l'asservissement, liée à une évolution différente des variables d'état du système réel et du modèle mathématique. Pour contrer cet inconvénient, la structure proposée est désormais modifiée pour faire intervenir explicitement la différence entre le système et le modèle. La philosophie est alors similaire à celle développée dans [Cutler et Ramaker, 1980] sous le nom de *DMC (Dynamic Matrix Control)*, et également par [Richalet *et al.*, 1987] via le formalisme de Différence Objet-Modèle (D.O.M.).

La différence  $\varepsilon_{s/m}$  entre la sortie du système  $A$  et celle du modèle  $A_{mod}$  est ainsi prise en compte selon le schéma de la Figure 5.4. La relation entre la sortie prédite du système  $\hat{A}$  et la sortie prédite du modèle  $\hat{A}_{mod}$  est la suivante :

$$\hat{A}_{k+j} = \hat{A}_{mod_{k+j}} + \underbrace{(A_k - A_{mod_k})}_{\varepsilon_{s/m}(k)}, \quad j = \overline{1, N} \quad (5.10.a)$$

Ce signal de D.O.M. n'intervient pas lorsque le modèle correspond parfaitement au système. Par ailleurs, il est possible d'aller encore plus loin et, de façon similaire à la notion d'auto-compensateur développée par [Richalet *et al.*, 1987], de faire intervenir l'accumulation de l'erreur après les  $j$  intervalles de prédiction par la relation :

$$\hat{A}_{k+j} = \hat{A}_{mod_{k+j}} + \underbrace{j(A_k - A_{mod_k})}_{\varepsilon_{s/m}(k)}, \quad j = \overline{1, N} \quad (5.10.b)$$



**Figure 5.4.** Principe de la structure de commande prédictive non-linéaire avec D.O.M.

Avec ce nouvel élément, le problème d'optimisation prend la forme suivante :

$$\min_{\chi} \sum_{j=1}^N (A_{set_{k+j}} - \hat{A}_{k+j})^2 + \eta \sum_{j=1}^N (F_{ref_{k+j-1}} - F_{in_{k+j-1}})^2$$

$$\text{sous les contraintes : } \begin{cases} \hat{A}_{mod_{k+1}} = H \mathcal{F}_{dis} f(x_{mod_k}, F_{in_k}) \\ \vdots \\ \hat{A}_{mod_{k+N}} = H \mathcal{F}_{dis} (\hat{x}_{mod_{k+N-1}}, F_{in_{k+N-1}}) \\ F_{in_k} \geq 0, \quad \forall k \in \mathcal{N} \\ x_k \geq 0, \quad \forall k \in \mathcal{N} \end{cases} \quad (5.11)$$

Dès lors, à chaque pas d'échantillonnage, la séquence optimale au sens du critère (5.8)  $\chi_{opt} = \{F_{in_{opt_k}}, \dots, F_{in_{opt_{k+N-1}}}, \hat{A}_{opt_{k+1}}, \dots, \hat{A}_{opt_{k+N}}\}$  est obtenue comme solution du problème d'optimisation (5.11). De cette séquence, seule la première commande  $F_{in_{opt_k}}$  est appliquée au bioprocédé à l'instant  $k$ . La même procédure est répétée à l'instant d'échantillonnage suivant selon le principe de l'horizon fuyant.

### 5.4.3 Résolution du problème d'optimisation

La résolution du problème d'optimisation sous contraintes (5.11) se heurte à deux difficultés majeures. La première difficulté réside dans la discrétisation du système (5.5) pour obtenir la formulation (5.7.c). En effet, la période d'échantillonnage doit être suffisamment faible pour que le modèle discret reste pertinent par rapport au modèle continu. Par rapport à la dynamique du système, elle sera alors beaucoup plus petite comparée au temps de réponse du système. Typiquement en effet, il convient de choisir  $T_e = 0,5 \text{ min}$  pour que le modèle discret reste pertinent (validation par comparaison des réponses indicielles par exemple), alors que le temps de réponse est plutôt de l'ordre de plusieurs dizaines de minutes, ce qui conduirait ensuite à des horizons de prédiction très importants et à une augmentation du nombre de variables de décision du problème d'optimisation.

La deuxième difficulté consiste en la présence de contraintes non-linéaires. Ceci augmente considérablement le temps de calcul de la boucle temps réel lors de la résolution du problème d'optimisation sous contraintes.

Pour contourner ces difficultés, il est alors possible de transformer le problème d'optimisation habituel précédent en un problème de programmation non-linéaire (PNL) et le résoudre en utilisant l'approche séquentielle CVP (*Control Vector Parametrization*). Cette stratégie déjà évoquée au chapitre 4 est développée dans [Ray, 1981] et [Vassiliadis, 1993] et se trouve utilisée de façon très classique pour les procédés chimiques et biochimiques, comme par exemple dans [Schlegel et Marquardt, 2006] et [Balsa-Canto *et al.*, 2001].

La méthode CVP se base sur la seule discrétisation des variables de commande, la séquence de commande étant approchée par une fonction constante par morceaux. L'avantage vient du fait qu'il n'est plus nécessaire de discrétiser le modèle continu, les états prédits étant obtenus par intégration du système différentiel continu (5.5). La période d'échantillonnage nécessaire au critère prédictif peut donc être choisie beaucoup plus grande que lors d'une discrétisation classique. La Figure 5.5 illustre cette stratégie.

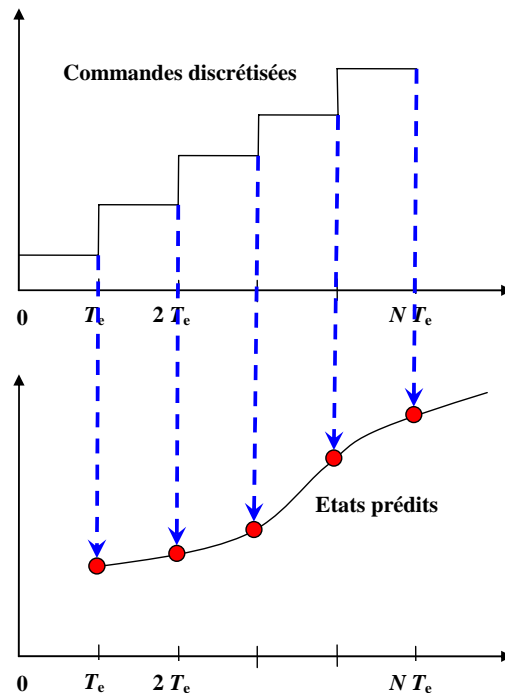


Figure 5.5. Principe de l'approche CVP.

Le problème d'optimisation (5.11) est ainsi transformé en un problème de programmation non-linéaire via la technique CVP. La nouvelle formulation du problème d'optimisation est alors :

$$\min_{\chi'} \sum_{j=1}^N (A_{set_{k+j}} - \hat{A}_{k+j})^2 + \eta \sum_{j=1}^N (F_{ref_{k+j-1}} - F_{in_{k+j-1}})^2 \quad (5.12)$$

sous les contraintes :  $F_{in_k} > 0, \forall k \in \mathcal{N}$

où  $\chi' = \{F_{in_k}, F_{in_{k+1}}, \dots, F_{in_{k+N-1}}\}$  est le nouveau vecteur d'optimisation. En effet, les contraintes liées au modèle et aux états disparaissent de ce vecteur puisqu'elles sont prises en compte d'une manière implicite lors de la prédiction des états via la méthode de la CVP. Ainsi, le problème d'optimisation (5.12) possède beaucoup moins de contraintes par rapport au problème initial (5.11). Si l'on souhaite mettre en œuvre un problème d'optimisation totalement non contraint, ce qui est toujours préférable lors d'une implantation temps réel, le changement de variable suivant :

$$F_{in} = \exp(v) \quad (5.13)$$

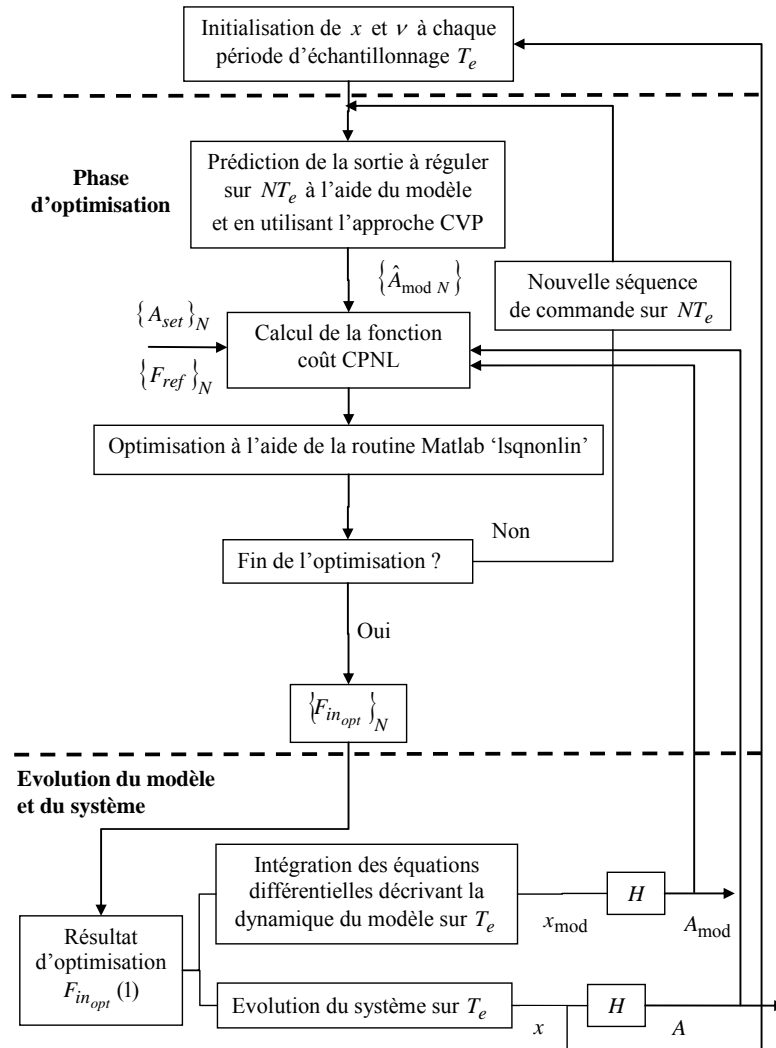
permet d'aboutir à la formulation du problème d'optimisation non contraint ci-dessous :

$$\min_{\chi''} \sum_{j=1}^N (A_{set_{k+j}} - \hat{A}_{k+j})^2 + \eta \sum_{j=1}^N (F_{ref_{k+j-1}} - \exp(v_{in_{k+j-1}}))^2 \quad (5.14)$$

où  $\chi'' = \{v_k, v_{k+1}, \dots, v_{k+N-1}\}$  est le vecteur d'optimisation.

Tenant compte de tous les développements précédents, la Figure 5.6 récapitule la structure globale proposée, qui inclut donc :

- un critère quadratique de commande prédictive non-linéaire (CPNL) incluant une différence objet-modèle,
- la transformation du problème CPNL en un problème de programmation non-linéaire résolu à l'aide de l'approche CVP,
- le changement de variable présentée par l'égalité (5.13).



**Figure 5.6.** Organigramme présentant la stratégie de commande proposée.

Lors des simulations proposées par la suite, la détermination de la séquence de commande optimale se fait en utilisant la fonction « *lsqnonlin* » de *Matlab*<sup>TM</sup>. Cette fonction permet de

résoudre les problèmes de moindres carrés non-linéaires en utilisant la méthode de *Gauss-Newton*.

## 5.5 Analyse des performances au nominal

La stratégie développée ci-dessus est maintenant implantée sous l'environnement *Matlab*<sup>TM</sup>. Rappelons que l'objectif de commande est de réguler la concentration en acétate à une valeur de référence  $A_{set}$  tout en forçant le débit d'alimentation à suivre une trajectoire de référence  $F_{ref}$ . Dans la suite de ce chapitre, on cherche essentiellement à illustrer les performances de la loi de commande proposée dans une configuration « nominale », à savoir en supposant que le modèle est parfait. Le signal D.O.M. introduit dans la prédiction n'agira donc pas. De même, les comparaisons avec d'autres techniques de commande se feront avec cette hypothèse de modèle adapté. Le chapitre 6 s'intéressera par la suite à tous les problèmes de robustesse et tiendra compte alors d'erreurs de modèle.

Pour illustrer les performances de la loi de commande développée au nominal, trois scénarii sont successivement envisagés. Le but du premier scénario est de réguler la concentration en acétate à la valeur  $A_{set} = 0$  g/kg, correspondant au fonctionnement biologique optimal. En pratique, à cause des problèmes de sensibilité des capteurs, ce premier scénario s'avère peu réaliste. Il est préférable en effet de réguler à une valeur non nulle mais malgré tout proche de 0. Ceci implique au préalable de déterminer le nouveau débit d'alimentation de référence pour cette valeur de référence non nulle, choisie égale à 0,5 g/kg. Enfin un troisième scénario, régulant initialement à 0,5 g/kg, fait intervenir un changement de consigne de façon à valider le comportement de la régulation en régime transitoire (même si ce type de scénario est peu réaliste pour le bioprocédé considéré).

### 5.5.1 Régulation de la concentration en acétate à $A_{set} = 0$ g/kg

Rappelons que le but est de maximiser la quantité en biomasse, et donc la croissance de la biomasse ; nous avons montré au chapitre 4 que cela se traduit par :

- la régulation de la concentration en acétate à  $A_{set} = 0$  g/kg,
- le maintien du système à la frontière des deux régimes :  $F_{ref} = u_{path}|_{A=0}$ .

Afin de valider cet objectif et par conséquent la stratégie de commande proposée, le premier scénario envisagé est le suivant :

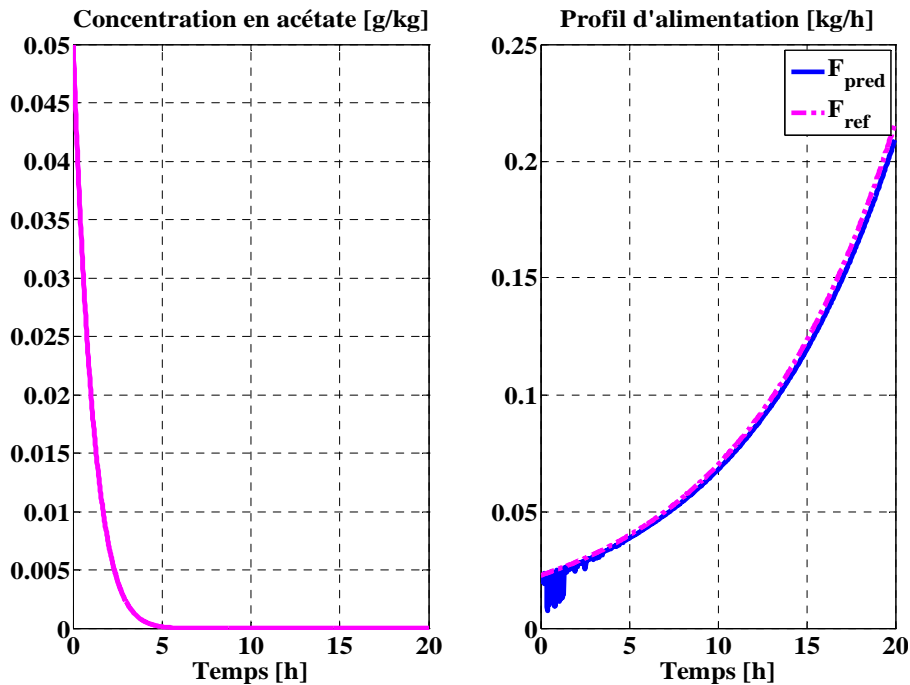
- l'objectif de commande est de réguler la concentration en acétate à  $A_{set} = 0$  g/kg tout en poursuivant le profil de référence  $F_{ref} = u_{path}|_{A=0}$ , relation (4.44.d),
- la période d'échantillonnage considérée est  $T_e = 2$  min. Elle est déterminée en fonction du temps de réponse du système, et des possibilités des différents capteurs permettant les mesures en ligne. Elle doit permettre également de satisfaire les impératifs liés à une éventuelle mise en œuvre de capteurs logiciels,

- la durée de simulation est choisie égale à 20 heures. Ceci correspond en pratique au temps moyen des cultures de laboratoire en mode fed-batch,
- les paramètres de commande choisis sont  $N = 8$  pour l'horizon de prédiction et  $\eta = 0,1$  pour le facteur de pondération sur la commande.

Ces dernières valeurs seront justifiées ultérieurement dans le paragraphe concernant l'étude de l'influence des paramètres de réglage du critère,

- les valeurs initiales des variables d'état du modèle sont  $x_0 = [X_0 \ S_0 \ A_0 \ W_0] = [5\text{g/kg} \ 0,03\text{g/kg} \ 0,05\text{g/kg} \ 3,17\text{kg}]$ .

La Figure 5.7 présente les évolutions de la concentration en acétate, du débit d'alimentation  $F_{pred}$  issu de la commande prédictive non-linéaire et du profil de référence  $F_{ref} = u_{path}|_{A=0}$ .

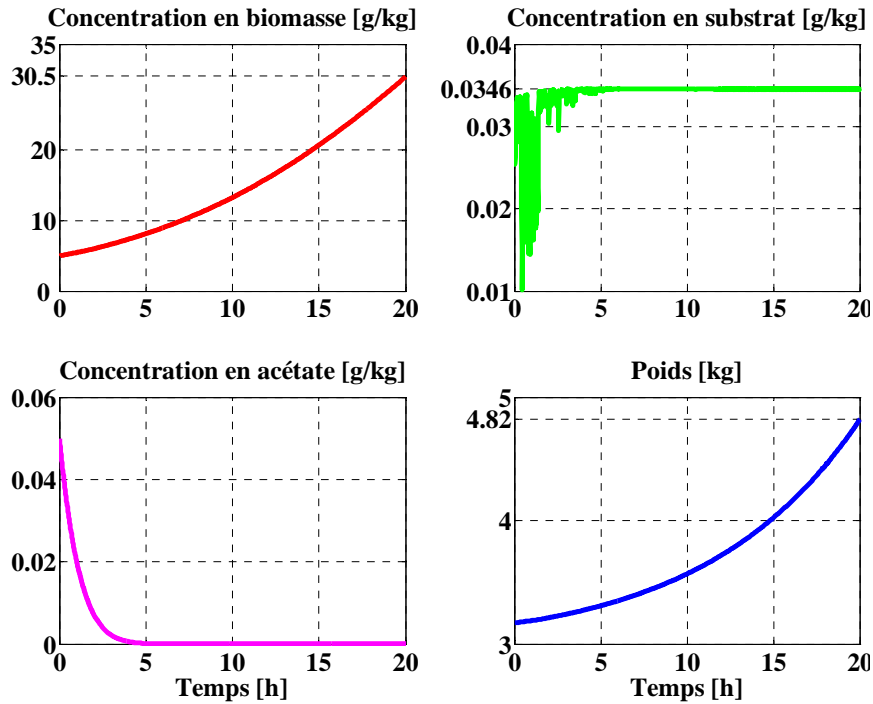


**Figure 5.7.** Evolution de la concentration en acétate, du débit d'alimentation et du profil de référence.

Cette figure montre que l'objectif de commande est atteint. En effet, la concentration en acétate est régulée à la valeur désirée  $A_{set} = 0 \text{ g/kg}$  et le débit d'alimentation suit le profil de référence.

La Figure 5.8 présente les évolutions des quatre variables d'état du système, détaillant ainsi l'effet de la commande prédictive sur les autres variables d'état. Elle montre que la concentration en biomasse a bien l'allure exponentielle attendue, résultat classiquement répandu chez les biologistes. On s'aperçoit par ailleurs que le fait d'avoir régulé la concentration en acétate induit implicitement une régulation de la concentration en substrat à la valeur  $0,0346 \text{ g/kg}$ , ce qui vérifie l'équation  $q_S = q_{Scrit}$  pour  $A = 0 \text{ g/kg}$ , et remplit donc la condition imposant de se trouver à la frontière entre les deux régimes de fonctionnement. On

voit apparaître l'idée que l'on régule en fait indirectement un couple  $(A, S)$  pour se trouver dans les conditions de fonctionnement optimales du bioréacteur. Enfin, la quantité de biomasse à l'instant final est de 147 g, ce qui constitue pour les conditions initiales considérées la quantité de biomasse produite la plus forte possible dès lors que le bioréacteur évolue à la frontière entre les deux régimes de fonctionnement.



**Figure 5.8.** Evolutions des variables d'état du modèle (5.5) en appliquant la commande prédictive non-linéaire.

## 5.5.2 Régulation de la concentration en acétate à $A_{set} = 0,5$ g/kg

La régulation de la concentration en acétate à une valeur nulle permet de maximiser la quantité de biomasse. Néanmoins, en pratique, les capteurs ont une certaine sensibilité. Comme il a été dit précédemment, ils ne permettent pas de mesurer le zéro précisément. Pour éviter ce problème, un deuxième scénario propose de réguler la concentration en acétate  $A$  à une valeur de référence  $A_{set}$  proche de 0 mais non nulle, ce qui constitue une solution sous-optimale et nécessite au préalable le calcul du profil associé.

### 5.5.2.1 Détermination du profil d'alimentation sous-optimal associé

En voulant réguler la concentration en acétate à une valeur non nulle, le profil de référence ne correspond plus au profil permettant de se maintenir sur la frontière des deux régimes. En effet, l'acétate est produit, le bioprocédé fonctionne alors en régime oxydo-fermentatif et la condition à la frontière ( $q_S = q_{Scrit}$ ) n'est plus valable. Il faut donc déterminer la trajectoire de référence  $F_{ref}$  correspondant à une valeur de référence  $A_{set} \neq 0$  g/kg.

Rappelons que l'équation différentielle de la concentration en acétate est :

$$\frac{dA}{dt} = r_A X - \frac{A}{W} F_{in} \quad (5.15.a)$$

où  $r_A = k_3 \frac{q_S - q_{Scrit}}{k_2}$  en régime oxydo-fermentatif (Tableau 5.1).

La valeur de référence  $A_{set}$  est constante, donc la dérivée  $dA/dt$  est nulle. D'où l'expression du profil de référence déduite de (5.15.a) :

$$F_{ref} = \frac{r_A}{A} XW \Big|_{A=A_{set}} \quad (5.15.b)$$

Pour maintenir le fonctionnement du bioprocédé en régime oxydo-fermentatif, la valeur de la concentration en substrat doit être également constante en accord avec la théorie du goulot d'étranglement et dès lors que l'on reste toujours dans le même régime. A partir de l'équation différentielle portant sur  $S$ , nous obtenons une expression du profil de référence  $F_{ref}$  vérifiant la condition  $dS/dt = 0$  :

$$\frac{dS}{dt} = -q_S X + \frac{(S_{in} - S)}{W} F_{in} \quad \text{soit} \quad F_{ref} = \frac{q_S}{(S_{in} - S)} XW \Big|_{S=cte} \quad (5.16)$$

A partir des deux expressions de  $F_{ref}$  (5.15.b) et (5.16), nous remarquons que la trajectoire de référence doit être calculée pour un couple  $(A_{set}, S_{crit})$  tel que :

$$\frac{q_S}{(S_{in} - S)} \Big|_{S=S_{crit}} = \frac{r_A}{A} \Big|_{A=A_{set}, S=S_{crit}} \quad (5.17)$$

Il s'avère que pour chaque valeur  $A_{set}$  est associée deux valeurs possibles de substrats  $S_{crit}$  (Annexe 3). Une seule de ces valeurs est physiquement acceptable. En effet, les deux solutions possibles sont soit une valeur faible en substrat, qui sera la valeur retenue ; ou bien une deuxième valeur proche de  $S_{in}$ , et qui ne sera pas prise en compte car elle correspond à des conditions opératoires non réalisables.

L'expression de la trajectoire de référence pour un couple  $(A_{set}, S_{crit})$  s'écrit alors sous la forme suivante :

$$F_{ref} = \frac{r_A}{A} XW \Big|_{A=A_{set}, S=S_{crit}} = \frac{q_S}{(S_{in} - S)} XW \Big|_{A=A_{set}, S=S_{crit}} \quad (5.18)$$

Notons que pour  $A_{set} = 0$ , nous retrouvons l'expression du profil optimal qui maintient la trajectoire sur la frontière des deux régimes  $u_{path}$  déterminé au chapitre 4 (équation (4.44.d)). L'expression (5.18) est donc un cas général déterminant le profil de référence  $F_{ref}$  pour un couple  $(A_{set}, S_{crit})$  donné, où la valeur de  $S_{crit}$  est calculée pour une valeur de référence  $A_{set}$  donnée à partir de l'équation (5.17). On met de nouveau en évidence cette notion de couple  $(A, S)$  déjà observée lors du scénario précédent.

Afin d'obtenir une relation temporelle issue de (5.18), reprenons les équations différentielles de la concentration en biomasse et du poids:

$$\begin{aligned} \frac{dX}{dt} &= r_X X - \frac{X}{W} F_{in} \\ \frac{dW}{dt} &= F_{in} \end{aligned} \quad (5.19)$$

A partir de ces deux équations, on obtient l'équation différentielle de la quantité de biomasse  $XW$  suivante :

$$\frac{d(XW)}{dt} = r_X XW \quad (5.20)$$

Le taux  $r_X$  dépend des concentrations en substrat et en acétate uniquement. Il est donc constant pour un couple  $(A_{set}, S_{crit})$  donné. Ainsi,  $XW$  solution de l'équation différentielle (5.20) est de la forme :

$$XW = X_0 W_0 e^{r_X t} \quad (5.21)$$

où  $X_0$  est la valeur initiale de la concentration en biomasse et  $W_0$  la valeur initiale du poids.

En utilisant les équations (5.18) et (5.21), l'expression du profil de référence  $F_{ref}$  s'écrit sous la forme finale :

$$F_{ref} = \frac{r_A}{A} X_0 W_0 e^{r_X t} \Big|_{A=A_{set}, S=S_{crit}} = \frac{q_S}{(S_{in} - S)} X_0 W_0 e^{r_X t} \Big|_{A=A_{set}, S=S_{crit}} \quad (5.22)$$

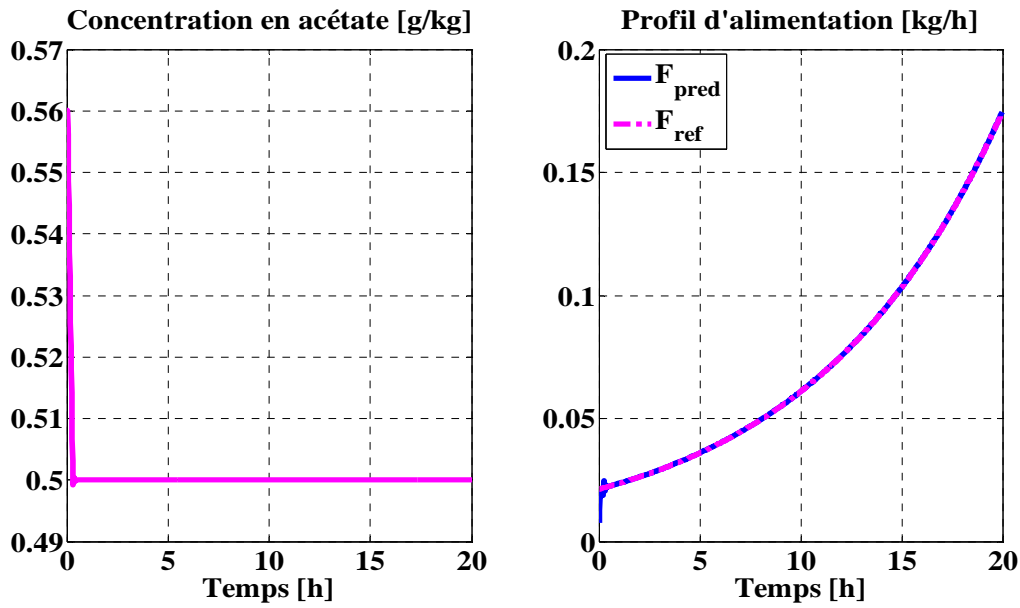
Cette expression montre que le profil de référence est un profil exponentiel, ce qui corrobore la connaissance et la pratique des biologistes à ce sujet. En effet, ceux-ci, reprenant l'allure (souvent exponentielle pour les bioprocédés) de l'évolution de la biomasse obtenue en phase batch, appliquent cette même allure au profil d'alimentation en phase fed-batch.

### 5.5.2.2 Résultats de simulation

Pour ce scénario, on se fixe comme objectif de réguler la concentration en acétate à la valeur  $A_{set} = 0,5$  g/kg tout en poursuivant la trajectoire de référence  $F_{ref}$  (5.22) calculée pour le couple  $(A_{set} = 0,5$  g/kg,  $S_{crit} = 0,0319$  g/kg). Rappelons que la valeur de  $S_{crit}$  est calculée à partir de l'équation (5.17) pour la valeur  $A_{set} = 0,5$  g/kg. Les conditions choisies pour la simulation sont :

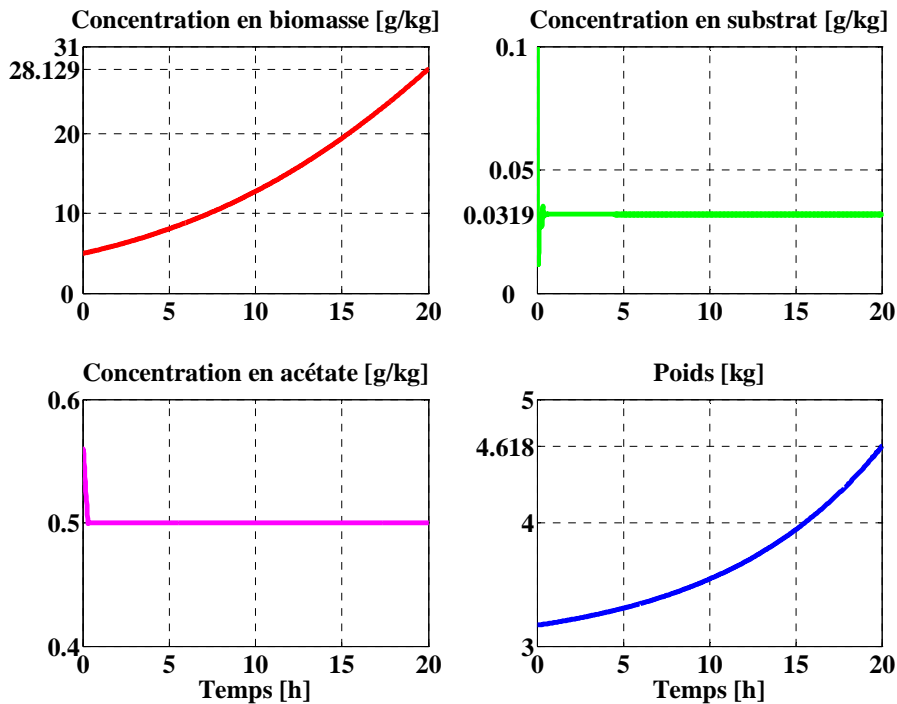
- la période d'échantillonnage est  $T_e = 2$  min,
- la durée de simulation est de 20 heures,
- l'horizon de prédiction est  $N = 8$ ,
- le facteur de pondération sur la commande  $\eta = 0,1$ ,
- le vecteur d'état initial est :  $x_0 = [X_0 \ S_0 \ A_0 \ W_0] = [5\text{g/kg} \ 0,1\text{g/kg} \ 0,55\text{g/kg} \ 3,17\text{kg}]$ .

La Figure 5.9 présente les évolutions de la concentration en acétate, du débit d'alimentation  $F_{pred}$  qui résulte de l'application de la commande prédictive non-linéaire et du profil de référence  $F_{ref}$ . On constate que la concentration en acétate est régulée à la valeur désirée  $A_{set} = 0,5$  g/kg et que le débit d'alimentation  $F_{pred}$  suit la trajectoire de référence  $F_{ref}$ . Ainsi, l'objectif de commande fixé est atteint.



**Figure 5.9.** Evolution de la concentration en acétate, du débit d'alimentation et du profil de référence sur 20 heures.

La Figure 5.10 présente les réponses des quatre variables d'état du système sur la durée de simulation. Elle montre que l'évolution de la concentration en biomasse a une forme exponentielle et que la quantité de biomasse à l'instant final est de 130 g, alors qu'elle est de 147 g quand on régule à  $A_{set} = 0$  g/kg. La solution sous-optimale proposée reste alors acceptable puisque la différence ne dépasse pas les 11%. Cette figure montre également que la concentration en substrat est régulée à la valeur  $S_{crit} = 0,0319$  g/kg. On remarque alors que le système converge, après une phase transitoire, vers le couple désiré  $(A_{set}, S_{crit})$ .



**Figure 5.10.** Evolutions des variables d'état du modèle (5.5) en appliquant la commande prédictive non-linéaire.

Finalement, en se fixant comme objectif de réguler la concentration en acétate avec la poursuite de la trajectoire de référence  $F_{ref}$ , on arrive à réguler la concentration en substrat à la valeur attendue.

Les Figures 5.11 et 5.12 présentent les évolutions des concentrations en acétate et en substrat respectivement sur la durée de simulation qui est de 20 heures et en zoomant sur l'intervalle [0 1 h].

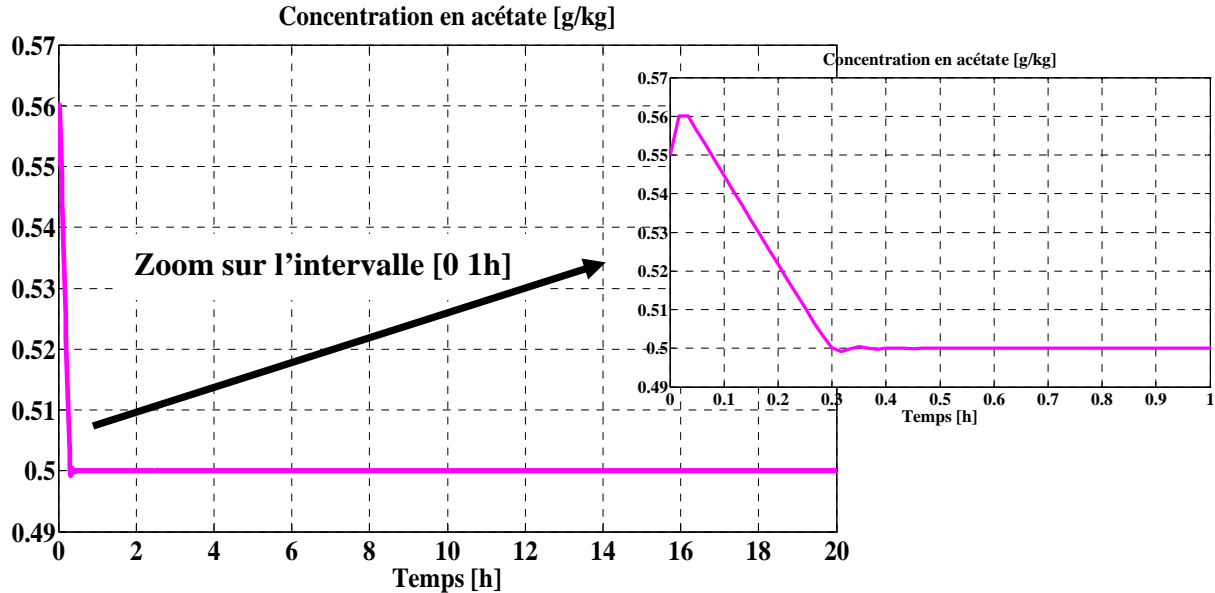


Figure 5.11. Evolution de la concentration en acétate sur 20 heures et zoom sur une heure.

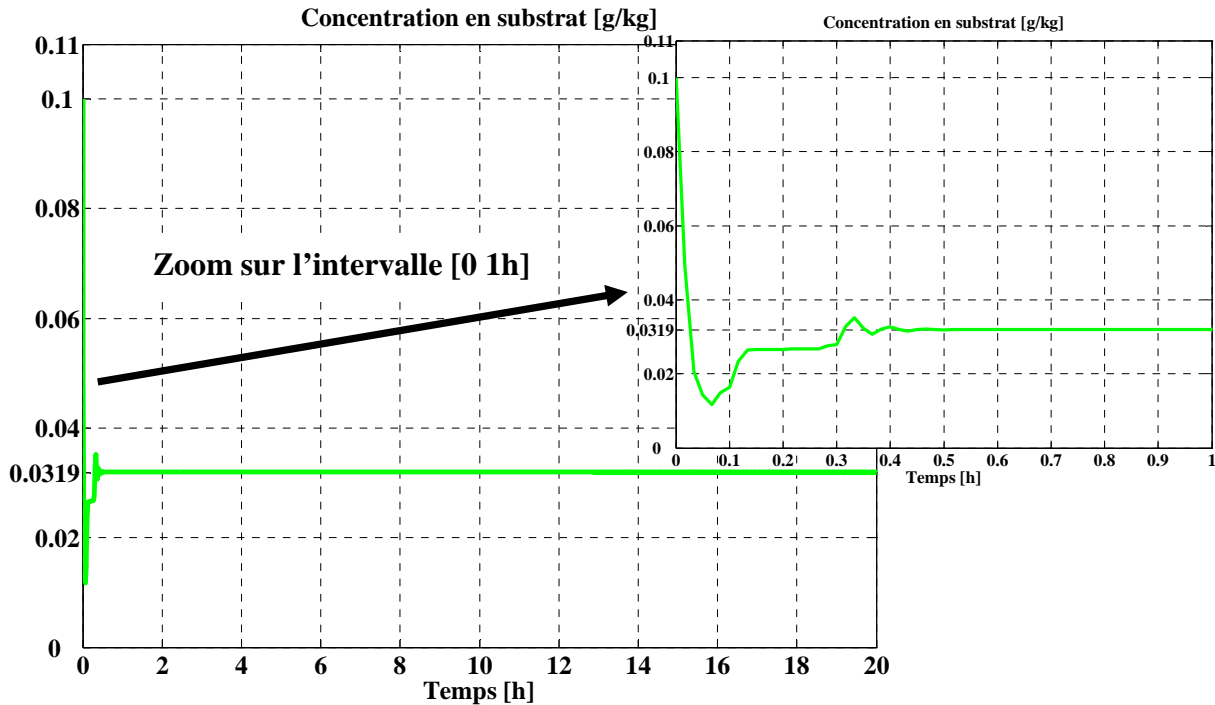
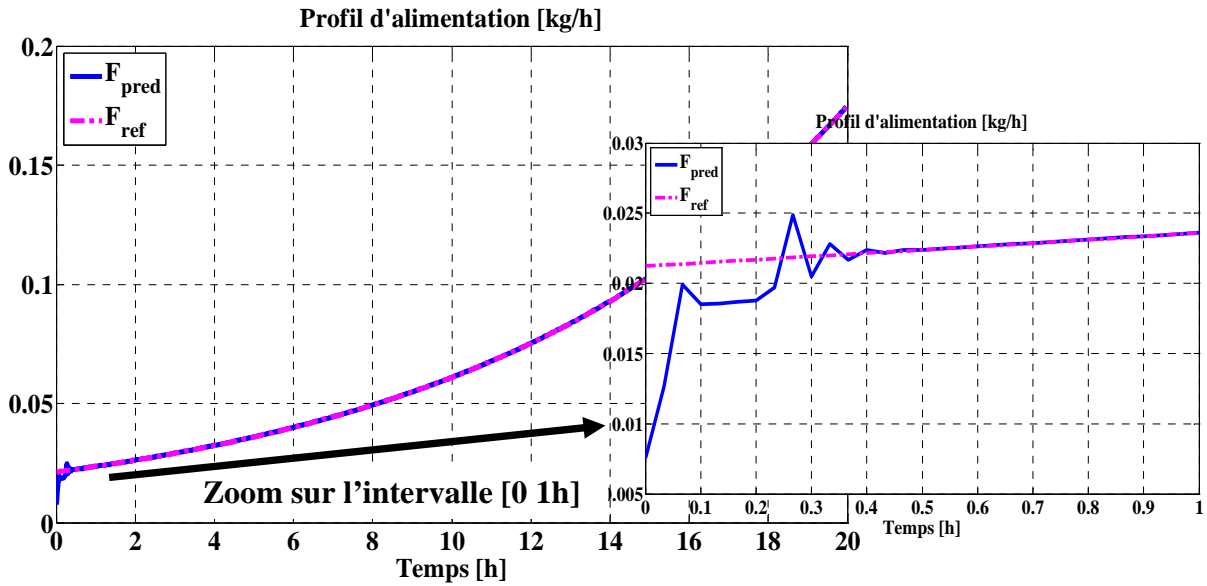


Figure 5.12. Evolution de la concentration en substrat sur 20 heures et zoom sur une heure.

Ces figures montrent que les concentrations en acétate et en substrat atteignent respectivement les valeurs désirées  $A_{set} = 0,5$  g/kg et  $S_{crit} = 0,0319$  g/kg au bout de 24 minutes, ce qui

caractérise la durée de la phase transitoire. Pour mieux l'illustrer, considérons également l'évolution du débit d'alimentation issu de la commande prédictive non-linéaire  $F_{pred}$  par rapport au profil de référence  $F_{ref}$  (Figure 5.13).



**Figure 5.13.** Evolution du débit d'alimentation  $F_{pred}$  et du profil de référence  $F_{ref}$  sur 20 heures et zoom sur une heure.

On constate effectivement que le débit d'alimentation passe par une phase transitoire qui dure 24 minutes avant de converger vers le profil de référence  $F_{ref}$  déterminé précédemment. Les Figures 5.11, 5.12 et 5.13 montrent qu'à partir de 24 minutes, le couple  $(A_{set}, S_{crit})$  est atteint et le débit d'alimentation suit la trajectoire de référence. Ceci assure le fonctionnement du bioprocédé en régime oxydo-fermentatif.

Mais que se passe-t-il pendant la phase transitoire ? Comment fonctionne le bioprocédé ? Rappelons que les taux  $r_X$  et  $r_A$  décrits dans le Tableau 5.1 s'écrivent également en fonction des taux spécifiques  $\mu_1$ ,  $\mu_2$  et  $\mu_3$  :

$$r_X = \mu_1 + \mu_2 + \mu_3 \quad \text{et} \quad r_A = k_3\mu_2 - k_4\mu_3 \quad (5.23)$$

Le Tableau 5.3 présente les expressions des taux spécifiques  $\mu_1$ ,  $\mu_2$  et  $\mu_3$  en fonction du régime de fonctionnement.

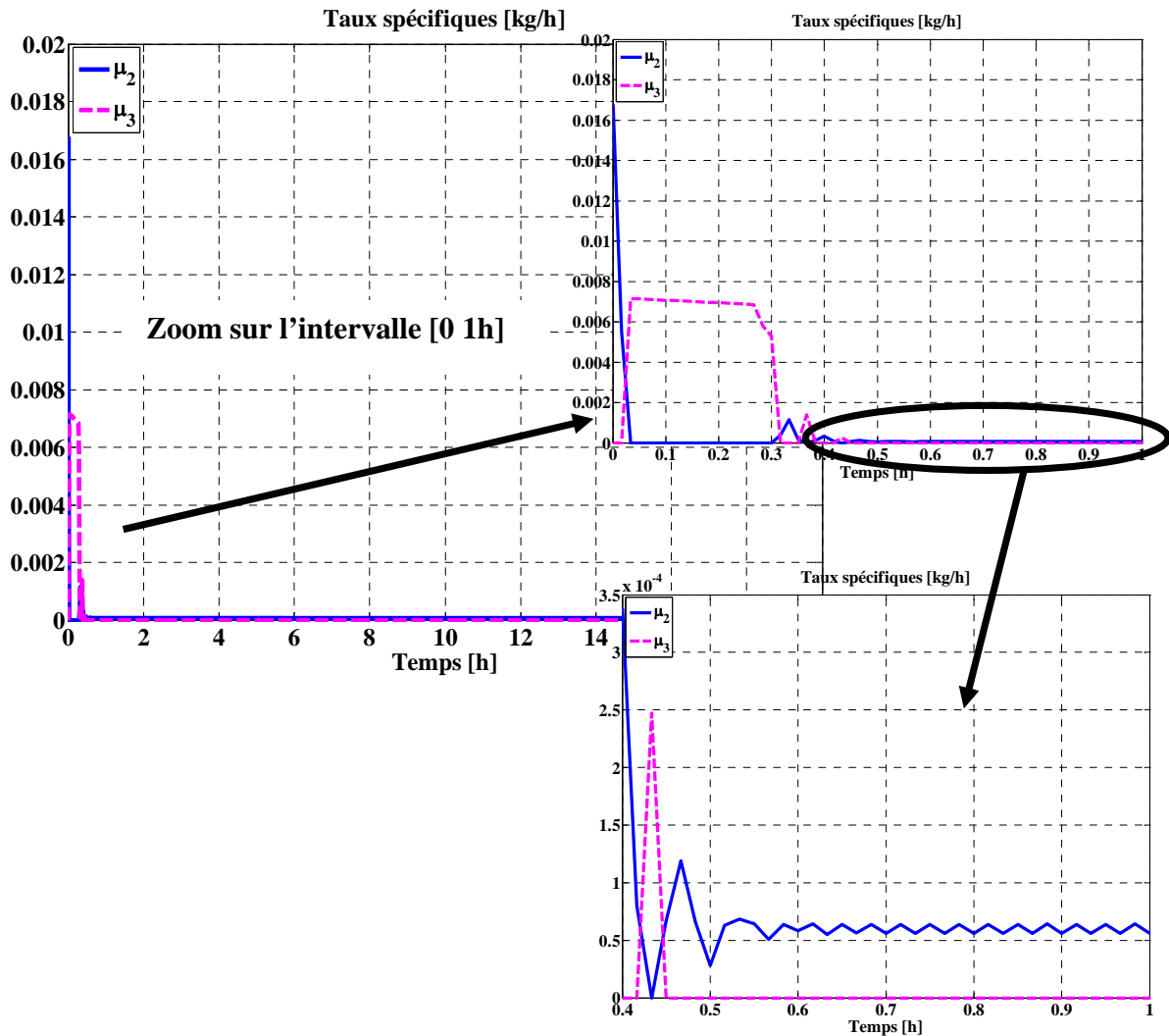
A partir de ce tableau, on remarque que les valeurs des taux spécifiques  $\mu_2$  et  $\mu_3$  peuvent être un outil qui nous permettra de déterminer les régimes de fonctionnement du bioprocédé pendant une durée de simulation donnée. En effet :

- $\mu_2 = 0$  correspond au fonctionnement en régime oxydatif,
- $\mu_3 = 0$  correspond au fonctionnement en régime oxydo-fermentatif.

Dans ce sens, la Figure 5.14 présente les évolutions des taux spécifiques  $\mu_2$  et  $\mu_3$  sur 20 heures en zoomant sur une heure puis sur l'intervalle [24 min 1h].

Type de régime	Expression de $\mu_1$	Expression de $\mu_2$	Expression de $\mu_3$
Régime oxydatif	Sous-régime 1	$\frac{q_S}{k_1}$	$\frac{q_{AC}}{k_4}$
	Sous-régime 2	$\frac{q_S}{k_1}$	$k_{os} \frac{q_{Scrit} - q_S}{k_{oa}}$
Régime oxydo-fermentatif	$\frac{q_{Scrit}}{k_1}$	$\frac{q_S - q_{Scrit}}{k_2}$	0

Tableau 5.3. Expressions des taux spécifiques selon le régime de fonctionnement.


 Figure 5.14. Evolutions des taux spécifiques  $\mu_2$  et  $\mu_3$  sur 20 heures avec zoom sur une heure et zoom sur l'intervalle [0,4h 1h].

Elle montre qu'il y a changement de régime durant la phase transitoire. En effet, le bioprocédé démarre en fonctionnant en régime oxydo-fermentatif, puis en régime oxydatif pendant environ 18 minutes, revient quelques minutes ensuite en régime oxydo-fermentatif puis en oxydatif et reste enfin en régime oxydo-fermentatif le tout environ 27 minutes après. Puisque les valeurs de  $\mu_2$  et de  $\mu_3$  sont faibles, le dernier zoom présenté permet de les distinguer. On constate clairement que le taux spécifique  $\mu_3$  est nul à partir de 27 minutes. Ce changement de régime explique le comportement du débit d'alimentation pendant la phase transitoire. Le fonctionnement du bioprocédé traverse les différents régimes avant de se maintenir au régime oxydo-fermentatif.

### 5.5.3 Régulation de la concentration en acétate avec changement de consigne

Même si ce dernier scénario n'a aucun intérêt pratique au niveau du protocole d'une culture, il permet en introduisant une variation de consigne de visualiser le comportement de la structure d'asservissement en régime transitoire. Ce comportement mis en évidence lors d'un changement de référence sera similaire en présence de perturbations éventuelles.

L'exemple de simulation présenté dans ce paragraphe reprend les mêmes conditions que les précédents :

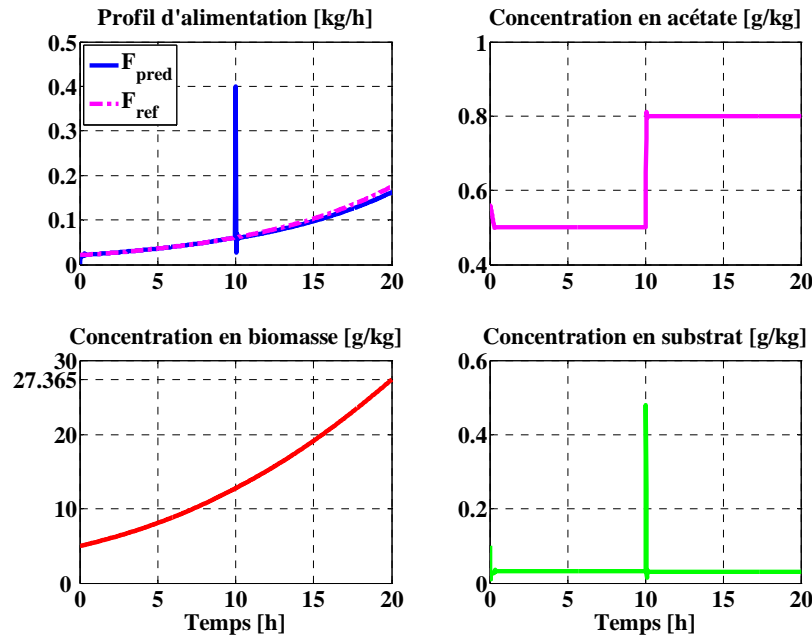
- la période d'échantillonnage est  $T_e = 2$  min ,
- la durée de simulation est de 20 heures,
- l'horizon de prédiction est  $N = 8$  ,
- le facteur de pondération sur la commande est  $\eta = 0,1$  ,
- les valeurs initiales des variables d'état :

$$[X_0 \quad S_0 \quad A_0 \quad W_0] = [5 \text{ g/kg} \quad 0,1 \text{ g/kg} \quad 0,55 \text{ g/kg} \quad 3,17 \text{ kg}]$$

La valeur désirée  $A_{set}$  est égale à 0,5 g/kg sur 10 heures puis atteint 0,8 g/kg ensuite. Le profil de référence considéré  $F_{ref}$  est le même que celui utilisé pour l'exemple précédent et calculé pour ( $A_{set} = 0,5$  g/kg,  $S_{crit} = 0,0319$  g/kg).

La Figure 5.15 montre les évolutions du profil d'alimentation, des concentrations en acétate, en biomasse et en substrat sur 20 heures pour un changement de consigne à 10 h. On remarque que la commande développée répond rapidement au changement de consigne. En effet, la concentration en acétate atteint la valeur désirée  $A_{set} = 0,5$  g/kg en moins d'une heure (similaire à l'exemple précédent), puis à partir de 10 h, elle atteint la valeur désirée  $A_{set} = 0,8$  g/kg encore plus rapidement.

L'évolution de la concentration en biomasse suit toujours une loi exponentielle. On observe un pic dans l'évolution du débit d'alimentation ainsi que dans celle de la concentration en substrat marquant le changement de consigne.

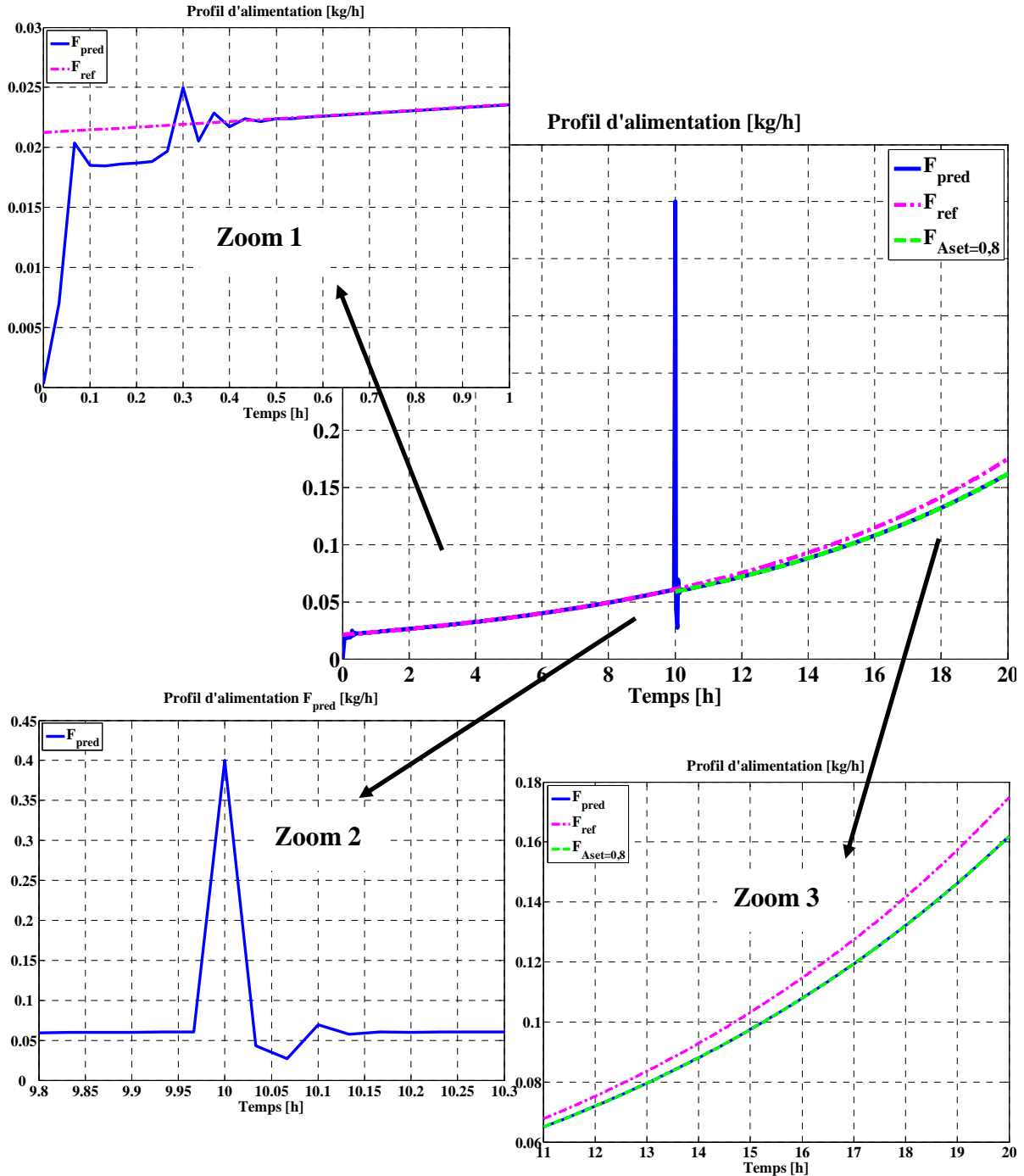


**Figure 5.15.** Evolution du profil d'alimentation, des concentrations en acétate, en biomasse et en substrat sur 20 heures.

Pour mieux analyser l'effet du changement de consigne sur l'évolution du débit d'alimentation issu de la commande prédictive non-linéaire, la Figure 5.16 présente cette évolution, celle de la trajectoire de référence  $F_{ref}$  qui rappelons-le, est calculée pour  $A_{set} = 0,5$  g/kg ainsi que le profil d'alimentation qui correspondrait à  $A_{set} = 0,8$  g/kg. Ce dernier est tracé uniquement pour faciliter l'interprétation des résultats. Trois zooms sont présentés pour mettre en évidence les informations que contient ce tracé.

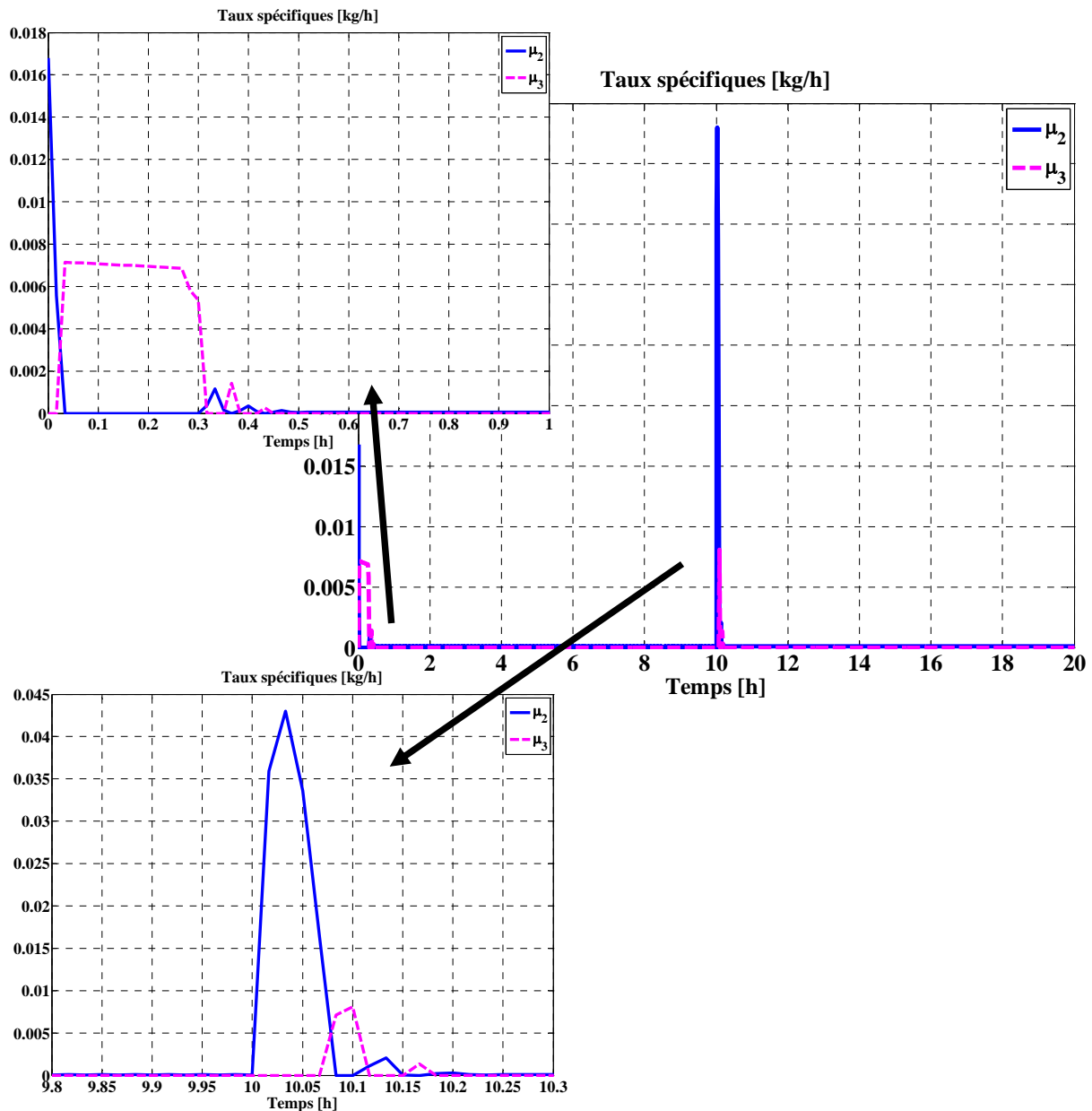
- sur l'intervalle [0 1 h] (Zoom1), l'évolution du débit d'alimentation est identique à celle présentée sur la Figure 5.13 car l'objectif à ce niveau est le même : réguler la concentration en acétate à  $A_{set} = 0,5$  g/kg avec poursuite de la trajectoire  $F_{ref}$  calculée pour le couple ( $A_{set} = 0,5$  g/kg,  $S_{crit} = 0,0319$  g/kg).
- à partir de 10 h, moment où nous avons décidé de changer de consigne (passage de la valeur  $A_{set} = 0,5$  g/kg à  $A_{set} = 0,8$  g/kg), on constate sur le Zoom 2 que ce changement implique une phase transitoire pendant laquelle le bioprocédé commute entre régime oxydatif et régime oxydo-fermentatif (voir Figure 5.17 pour l'évolution des taux spécifiques) avant de suivre une loi exponentielle tracée dans le Zoom 3.
- après cette phase de transition, la concentration en acétate est régulée à la nouvelle valeur  $A_{set} = 0,8$  g/kg comme le montre la Figure 5.15. On voit sur le Zoom 3 que le débit d'alimentation ne suit plus la trajectoire de référence  $F_{ref}$ . Ce résultat peut se comprendre dès lors que celle-ci est calculée pour le couple ( $A_{set} = 0,5$  g/kg,  $S_{crit} = 0,0319$  g/kg). On peut alors légitimement penser que le débit d'alimentation suit la trajectoire qui serait calculée pour  $A_{set} = 0,8$  g/kg et la valeur de concentration en substrat correspondante. C'est pour cette raison que nous avons ajouté le tracé de cette trajectoire sur la Figure 5.16 (Zoom 3), elle est notée  $F_{A_{set}=0,8}$  et est bien sûr initialisée avec les conditions initiales au moment du changement de

consigne. On voit que le débit d'alimentation suit effectivement cette nouvelle trajectoire de référence, sans tenir compte lors de l'optimisation des écarts dans le critère entre la commande et l'ancienne commande de référence. Ainsi, l'algorithme d'optimisation privilégie avant tout la régulation de la concentration en acétate, et la pondération sur la commande étant faible, le terme sur l'écart à la commande de référence n'intervient pas (ou très peu).



**Figure 5.16.** Evolution du débit d'alimentation, du profil de référence et du profil pour  $A_{set} = 0,8$  g/kg sur 20 heures avec 3 zooms.

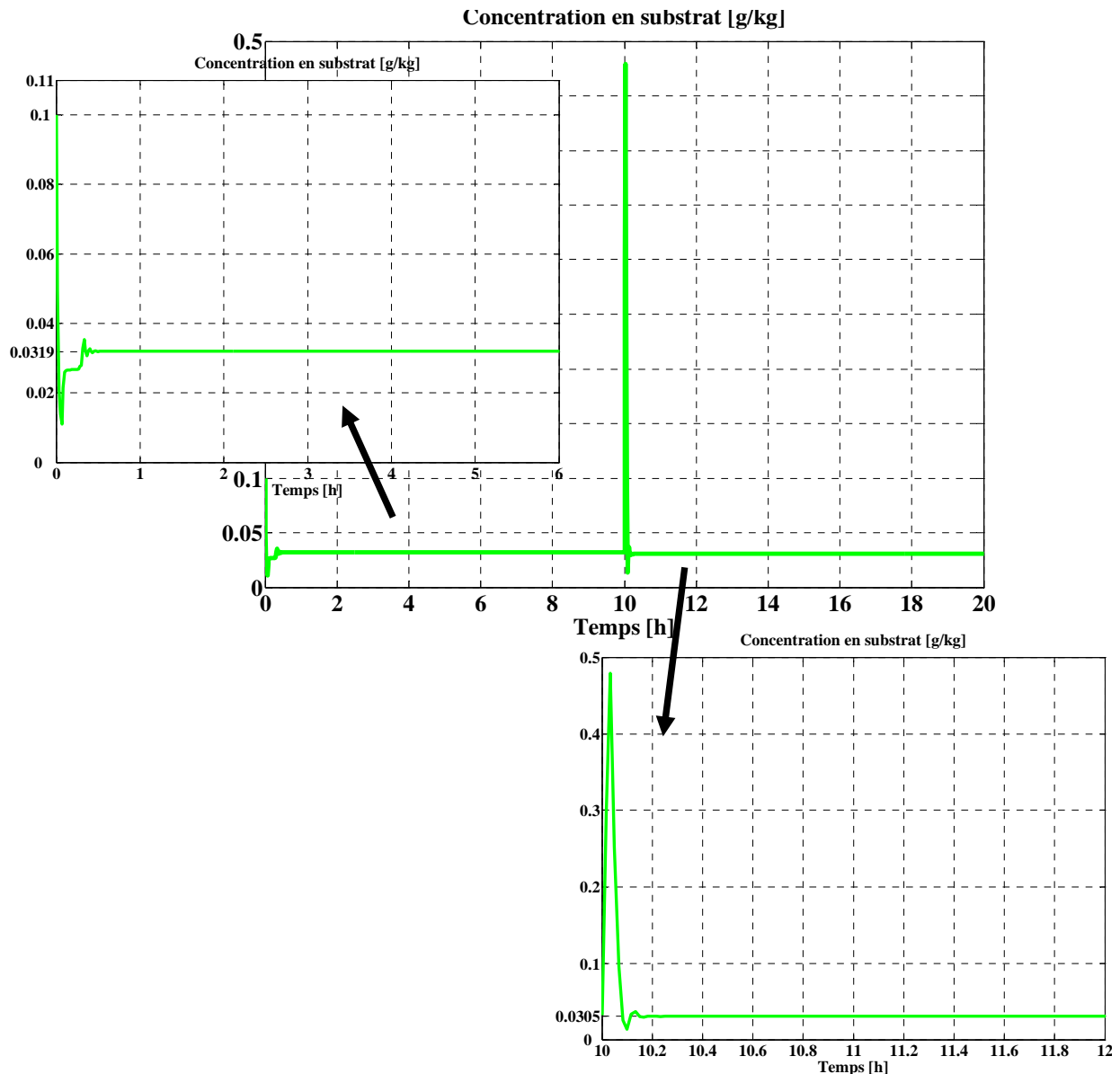
La Figure 5.17 présente les évolutions des taux spécifiques  $\mu_2$  et  $\mu_3$  qui déterminent le régime de fonctionnement du bioprocédé comme nous l'avons montré précédemment. Deux zooms sont également présentés pour une meilleure lecture du graphe. Sur cette figure, on remarque que la première phase transitoire correspond au même changement de régime que celui détaillé Figure 5.14. La deuxième phase transitoire qui nous intéresse ici commence au moment où l'on applique un changement de consigne. On remarque sur le deuxième zoom que le taux spécifique  $\mu_3$  devient non nul pendant quelques minutes. Ceci montre que le changement de consigne induit un changement de régime pendant quelques minutes avant de maintenir le fonctionnement du bioprocédé en régime oxydo-fermentatif.



**Figure 5.17.** Evolutions des taux spécifiques  $\mu_2$  et  $\mu_3$  sur 20 heures et zooms.

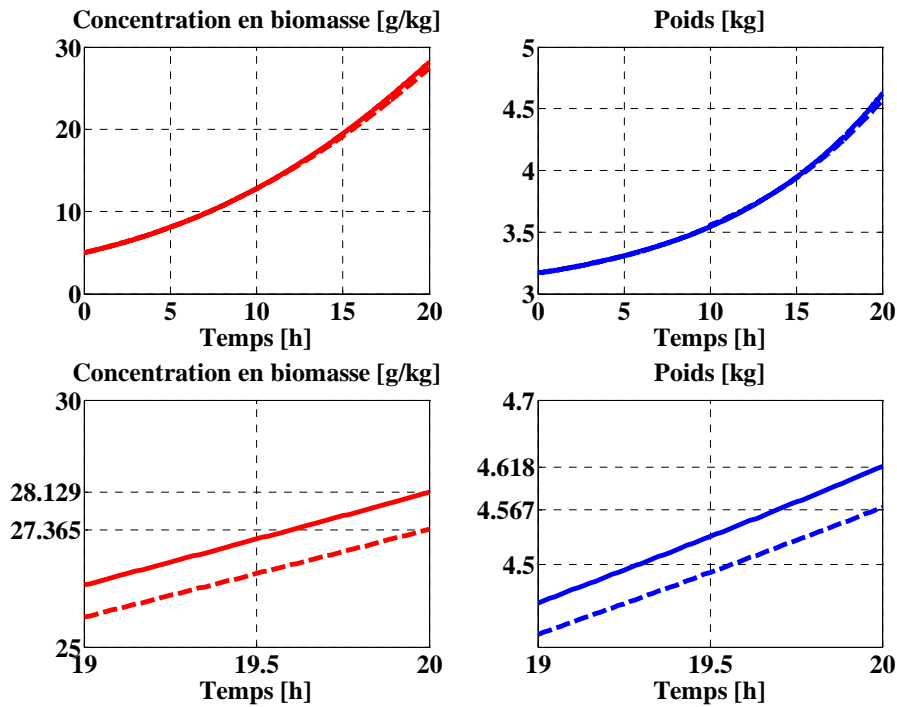
La Figure 5.15 a montré que la concentration en substrat atteint une valeur constante en même temps que la concentration en acétate. Nous présentons la Figure 5.18 pour analyser les valeurs atteintes par la concentration en substrat et leur signification. Avant le changement de

consigne, la concentration en acétate est régulée à la valeur désirée  $A_{set} = 0,5 \text{ g/kg}$ . On voit sur la Figure 5.18, que la concentration en substrat est régulée à la valeur  $0,0319 \text{ g/kg}$ , correspondante à  $S_{crit}$  et vérifiant l'équation (5.17). De la même manière, quand la concentration en acétate est régulée à  $A_{set} = 0,8 \text{ g/kg}$ , la concentration en substrat est régulée à la valeur  $0,0305 \text{ g/kg}$  qui vérifie aussi l'équation (5.17). D'autres essais ont été réalisés et ne sont pas tous présentés ici, mais permettent d'aboutir au même constat : quand la concentration en acétate est régulée à une valeur désirée  $A_{set}$ , la concentration en substrat est régulée à la valeur  $S_{crit}$  vérifiant l'équation (5.17). La régulation de la concentration en acétate induit donc la régulation implicite du couple  $(A, S)$ .



**Figure 5.18.** Evolution de la concentration en substrat sur 20 heures et zooms.

Pour terminer l'étude de ces scénarii, la Figure 5.19 présente les évolutions de la concentration en acétate et du poids pour les deux derniers exemples étudiés : régulation à une valeur constante et régulation avec changement de consigne. Cette comparaison a pour but de quantifier l'impact d'un changement de consigne sur la maximisation de la quantité  $XW$ .



**Figure 5.19.** Evolution de la concentration en acétate et du poids sur 20 heures avec zoom sur l'intervalle [19 h 20 h]. En trait continu, réponses sans changement de consigne et en pointillé, réponses avec changement de consigne.

Rappelons que l'objectif final de la commande développée est de maximiser la quantité  $XW$  à l'instant final. On voit sur cette figure que le premier exemple (régulation à  $A_{set} = 0,5$  g/kg) fournit une quantité de biomasse plus importante que le deuxième exemple (régulation avec changement de consigne), respectivement 129,9 g et 125 g. Ce résultat s'explique aussi d'un point de vue biologique. En effet, la valeur  $A_{set} = 0,8$  g/kg se place plus loin des conditions optimales comparativement à la configuration à  $A_{set} = 0,5$  g/kg, donc induit une perte d'optimalité plus grande.

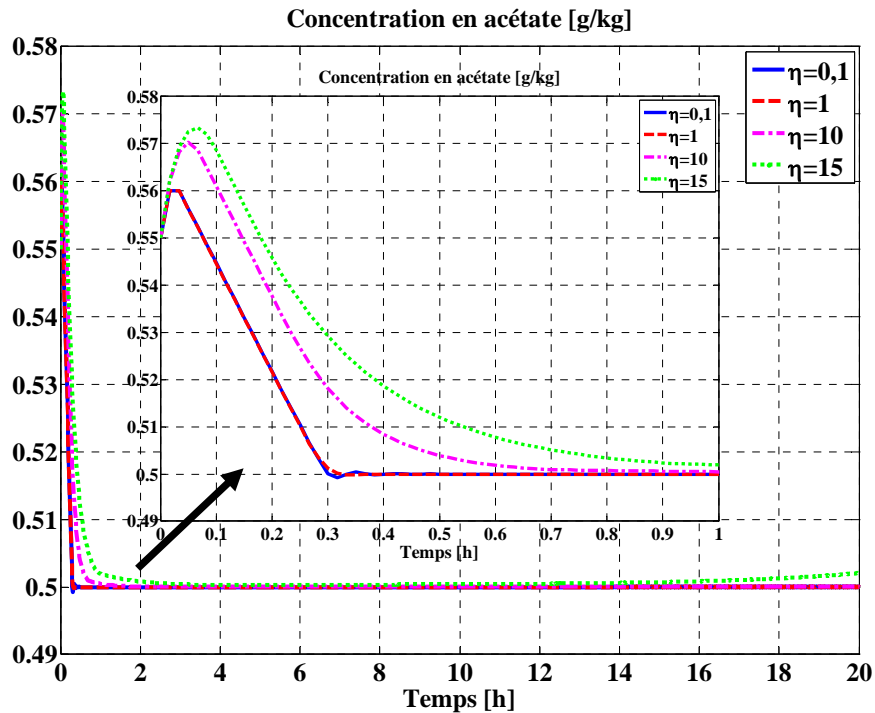
### 5.5.4 Influence des paramètres de commande

Ce paragraphe se focalise plus spécifiquement sur l'analyse de l'influence des paramètres de commande, à savoir le facteur de pondération sur la commande  $\eta$  et l'horizon de prédiction  $N$ . Comme nous l'avons précisé dans les deux exemples précédents, nous avons choisi comme valeurs de ces paramètres :

- $\eta = 0,1$  pour le facteur de pondération sur la commande,
- $N = 8$  pour l'horizon de prédiction.

Ce choix sera justifié à partir de l'analyse de quelques résultats de simulation présentés sur les figures suivantes. Pour tous ces tests, les conditions de simulation sont identiques à celles du deuxième exemple (régulation de la concentration en acétate à la valeur désirée  $A_{set} = 0,5$  g/kg avec poursuite de la trajectoire  $F_{ref}$ ), en faisant varier les valeurs des paramètres du critère  $\eta$  et  $N$ .

La Figure 5.20 présente l'évolution de la concentration en acétate pour différentes valeurs du facteur de pondération  $\eta$  et pour un horizon de prédiction  $N = 8$ . On remarque que le temps de réponse augmente quand la valeur de  $\eta$  augmente. On voit également que pour  $\eta = 15$ , la concentration en acétate n'est pas tout à fait régulée, du fait de la pondération très importante appliquée sur la commande. Ainsi, pour une régulation rapide de la concentration en acétate, la valeur du facteur de pondération doit être choisie faible, dans notre cas avec des valeurs inférieures à 1. On retrouve par ailleurs des résultats classiques en commande prédictive, montrant qu'une augmentation importante du facteur de pondération pénalise le temps de réponse du système asservi.



**Figure 5.20.** Evolution de la concentration en acétate pour  $N = 8$  et différentes valeurs de  $\eta$ .

A partir de l'analyse de la Figure 5.20, deux valeurs du facteur de pondération semblent raisonnables parmi celles testées. Pour pouvoir choisir une valeur qui nous servira pour les simulations à venir, nous utilisons la Figure 5.21 qui présente l'écart entre le débit d'alimentation issu de la commande prédictive  $F_{pred}$  et la trajectoire de référence  $F_{ref}$  pour différentes valeurs de  $\eta$ . Cette figure nous permet de déterminer la valeur pour laquelle la poursuite de la trajectoire est meilleure. On voit à l'aide du zoom que la différence  $F_{pred} - F_{ref}$  est minimale pour la valeur du facteur de pondération  $\eta = 0,1$ . Les résultats des Figures 5.20 et 5.21 justifient ainsi le choix du facteur de pondération  $\eta = 0,1$ .

Pour montrer l'influence de l'horizon de prédiction, la Figure 5.22 propose l'évolution du temps de calcul en minutes en fonction des valeurs de  $N$  et pour différentes valeurs du facteur de pondération  $\eta$ . On constate que le temps de calcul augmente lorsque l'horizon de prédiction s'accroît. Ce résultat est classique car à chaque pas d'échantillonnage le modèle de prédiction doit être intégré sur une durée  $N \times T_e$ , d'une part, et d'autre part le nombre de variables d'optimisation augmente. Globalement, un compromis doit être recherché, car, comme on vient de le voir, la valeur choisie pour l'horizon de prédiction ne doit pas être trop

grande, mais d'un autre côté, elle ne peut être trop petite non plus afin de fournir à l'algorithme une vision du futur suffisante pour atteindre les objectifs souhaités. Ainsi, dans notre cas, la valeur  $N = 8$  semble gérer au mieux ce compromis.

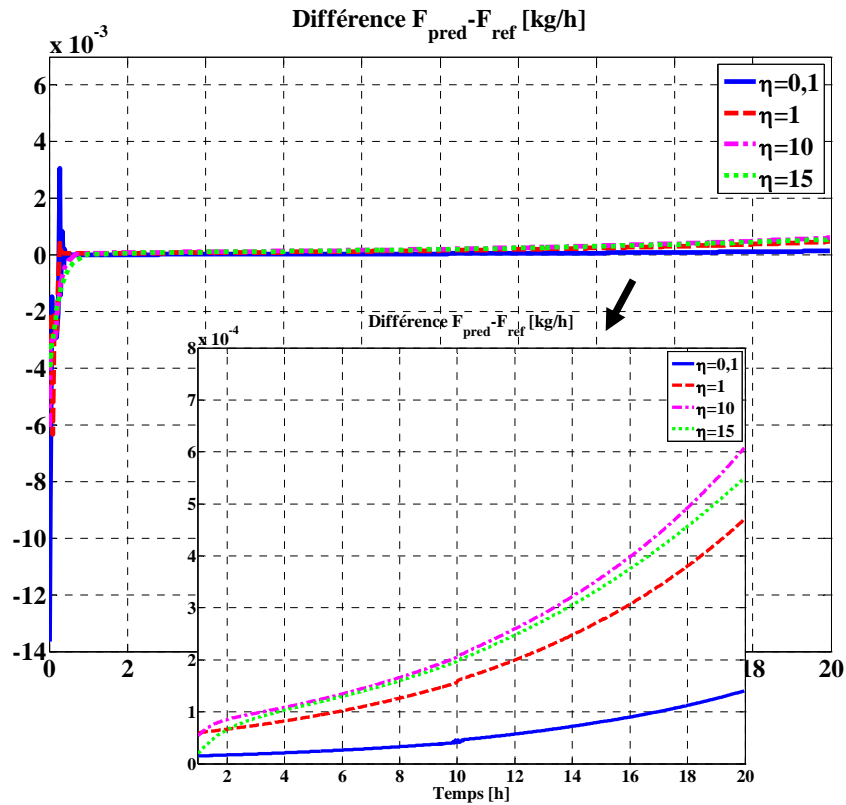


Figure 5.21. Tracé de l'écart  $F_{pred} - F_{ref}$  pour  $N = 8$  et différentes valeurs de  $\eta$ .

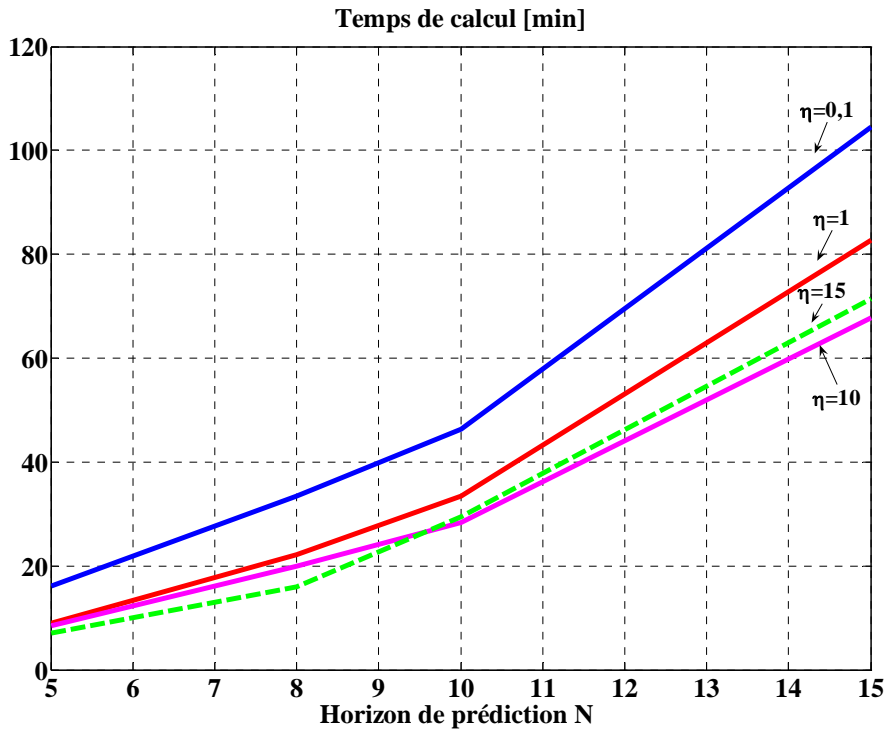


Figure 5.22. Temps de calcul en fonction de l'horizon de prédiction pour différentes valeurs du facteur de pondération.

**Remarques :**

- La valeur  $N = 5$  donne également de bons résultats au nominal. Mais, elle n'est pas choisie car en présence des modèles désadaptés lors de l'étude de robustesse (chapitre 6), il s'avère nécessaire de prédire sur un intervalle suffisamment grand pour une meilleure régulation.
- Les résultats non reproduits ici montrent par ailleurs qu'augmenter l'horizon de prédiction au-delà d'une certaine valeur (8 ici) n'apporte rien de plus en termes de dépassement, temps de réponse ... et augmente la complexité du problème.

L'analyse de ces trois figures procure une idée de l'influence de ces deux paramètres. Des règles relativement fines sont malgré tout difficiles à élaborer compte-tenu de leur interaction mutuelle et de la complexité de l'algorithme d'optimisation.

## 5.6 Comparaisons avec d'autres commandes

Les scénarii précédents ont permis de valider l'approche proposée et de dégager un certain nombre de performances intéressantes, dans le cas nominal. Afin de préciser maintenant l'utilité et l'opportunité de cette stratégie, une étude comparative est développée ci-dessous, dans le cas nominal toujours, mettant en jeu deux autres stratégies de commande rencontrées dans ce domaine : la commande adaptative linéarisante et la commande par modèle générique.

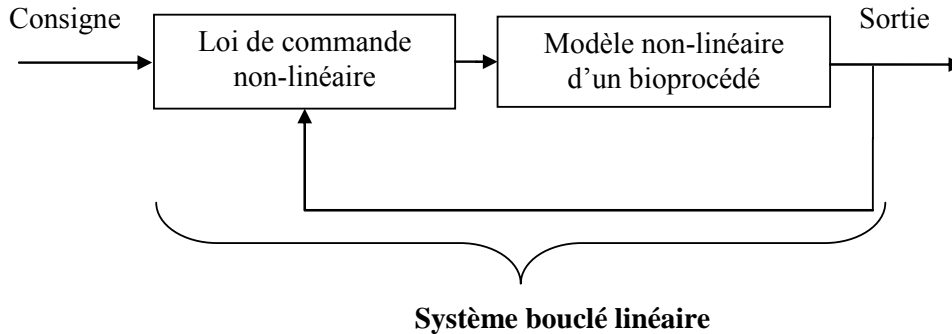
La littérature présente un seul exemple de régulation de la concentration en acétate pour le bioprocédé *E. coli* [Rocha, 2003] dans lequel est développée une commande adaptative linéarisante. L'application de celle-ci pour la régulation des bioprocédés est largement étudiée par Bastin et Dochain depuis les années 90, notamment dans [Bastin et Dochain, 1990]. Nous allons donc comparer notre stratégie de commande à celle appliquée au même bioprocédé dans [Rocha, 2003].

Par ailleurs, un régulateur par modèle générique GMC a également été développé sur notre bioprocédé afin de le comparer au régulateur prédictif élaboré. La commande GMC a été développée par Lee et Sullivan à la fin des années 80 [Lee et Sullivan, 1988]. C'est une technique de commande avancée qui prend en compte les non-linéarités du modèle dans la loi de commande. Elle est utilisée dans plusieurs applications industrielles. [Jenzsch *et al.*, 2006] présentent des résultats satisfaisants de l'application de cette commande pour la régulation du taux spécifique de croissance du bioprocédé de culture des *E. coli*. Toujours dans le domaine des biotechnologies, la commande GMC est également appliquée pour la digestion anaérobie [Costello *et al.*, 1989]. Dans [Cott *et al.*, 1989] et [Douglas *et al.*, 1994], elle est appliquée pour la régulation des colonnes à distiller.

### 5.6.1 Comparaison avec la commande adaptative linéarisante

La commande adaptative linéarisante est appliquée aux bioprocédés depuis les années 90 grâce à Bastin et Dochain [Bastin et Dochain, 1990]. Le principe de fonctionnement de cette commande est présenté Figure 5.23. L'objectif est d'appliquer une commande adaptative au modèle non-linéaire d'un bioprocédé afin d'obtenir un système en boucle fermée linéaire.

Comme mentionné précédemment, cette commande est développée et appliquée dans [Rocha, 2003] à la commande du bioprocédé de culture des bactéries *E. coli*. Nous allons utiliser la commande adaptative linéarisante telle qu'elle a été utilisée par Rocha pour la comparer à notre commande.



**Figure 5.23.** Principe de fonctionnement d'une commande linéarisante.

### 5.6.1.1 Commande adaptative linéarisante appliquée au bioprocédé *E. coli*

Nous présentons ici brièvement le développement de la commande adaptative linéarisante proposée dans [Rocha, 2003]. La commande développée est appliquée au modèle de simulation présenté au chapitre 2. Il est constitué de six équations différentielles décrivant les évolutions des six variables d'état considérées à savoir  $X$ ,  $S$ ,  $A$ ,  $O$ ,  $C$  et  $W$  qui représentent respectivement les concentrations en biomasse, substrat, acétate, oxygène et dioxyde de carbone et le poids.

Pour le développement de la loi de commande, il est proposé une réduction de modèle sous l'hypothèse que les dynamiques du substrat, de l'oxygène et du dioxyde de carbone sont rapides par rapport à celles de la biomasse et de l'acétate.

Le système d'équations différentielles décrivant les évolutions des variables d'état  $S$ ,  $O$  et  $C$  est :

$$\frac{d}{dt} \underbrace{\begin{pmatrix} S \\ O \\ C \end{pmatrix}}_{\xi_f} = \underbrace{\begin{bmatrix} -k_1 & -k_2 & 0 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix}}_{K_f} \underbrace{\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}}_r - D \underbrace{\begin{pmatrix} S \\ O \\ C \end{pmatrix}}_{\xi_f} + \underbrace{\begin{pmatrix} DS_{in} \\ OTR \\ -CTR \end{pmatrix}}_{F_f + Q_f} \quad (5.24)$$

où :

- $K_f$  est la matrice des coefficients stœchiométriques,
- $r$  est le vecteur des vitesses de réaction,
- $D$  est le taux de dilution,
- $S_{in}$  est la concentration en substrat dans le milieu d'alimentation,
- $OTR$  est le taux de transfert d'oxygène de la phase gazeuse vers le milieu de culture,
- $CTR$  est le taux de transfert du dioxyde de carbone du milieu de culture vers la phase gazeuse.

Les dynamiques de ces trois variables d'état sont négligées, et en écrivant  $\xi_f = \dot{\xi}_f = 0$ , on a :

$$K_f r + F_f + Q_f = 0 \quad (5.25)$$

Puisque la matrice  $K_f$  est de rang plein, le vecteur  $r$  est exprimé comme suit :

$$r = K_f^{-1}(-Q_f - F_f) \quad (5.26)$$

D'un autre côté, le système d'équations différentielles décrivant les évolutions des variables d'état  $X$  et  $A$  est donné par :

$$\frac{d}{dt} \underbrace{\begin{pmatrix} X \\ A \end{pmatrix}}_{\xi_s} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & k_3 & -k_4 \end{bmatrix}}_{K_s} \underbrace{\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}}_r - D \underbrace{\begin{pmatrix} X \\ A \end{pmatrix}}_{\xi_s} \quad (5.27)$$

En remplaçant le vecteur  $r$  dans l'équation (5.27) par son expression (5.26), il vient :

$$\frac{d\xi_s}{dt} = K_s r - D \xi_s \Rightarrow \frac{d\xi_s}{dt} = K_s K_f^{-1}(-Q_f - F_f) - D \xi_s \quad (5.28)$$

ainsi :

$$\frac{dA}{dt} = -p_1 CTR - p_2 OTR + p_3 D S_{in} - D A \quad (5.29)$$

où :

- $p_1$ ,  $p_2$  et  $p_3$  sont des paramètres exprimés en fonction des coefficients stœchiométriques  $k_j$  pour  $j \in [1,10]$ ,
- Les transferts  $OTR$  et  $CTR$  s'expriment en fonction des vitesses de réaction comme suit :  $OTR = k_5 r_1 + k_6 r_2 + k_7 r_3$  et  $CTR = k_8 r_1 + k_9 r_2 + k_{10} r_3$ . Ces expressions sont valables uniquement sous l'hypothèse citée ci-dessus.

Pour réguler la concentration en acétate à une valeur  $A_{set}$  donnée, une dynamique du premier ordre est imposée comme suit :

$$\frac{d}{dt}(A_{set} - A) = -\gamma(A_{set} - A) \quad \text{avec : } \gamma > 0 \quad (5.30)$$

où  $\gamma$  représente le gain de commande. Il s'exprime en fonction du temps de réponse à 5% désiré :  $\gamma = 3/Tr_{5\%}$ .

La loi de commande adaptative linéarisante s'écrit alors de la façon suivante :

$$D = \frac{p_1 CTR + p_2 OTR + \gamma(A_{set} - A)}{p_3 S_{in} - A} \quad (5.31)$$

On remarque que la variable de commande considérée est le taux de dilution. On peut en déduire le débit d'alimentation à l'aide de la relation :

$$F_{in} = DW \quad (5.32)$$

### 5.6.1.2 Résultats de simulation

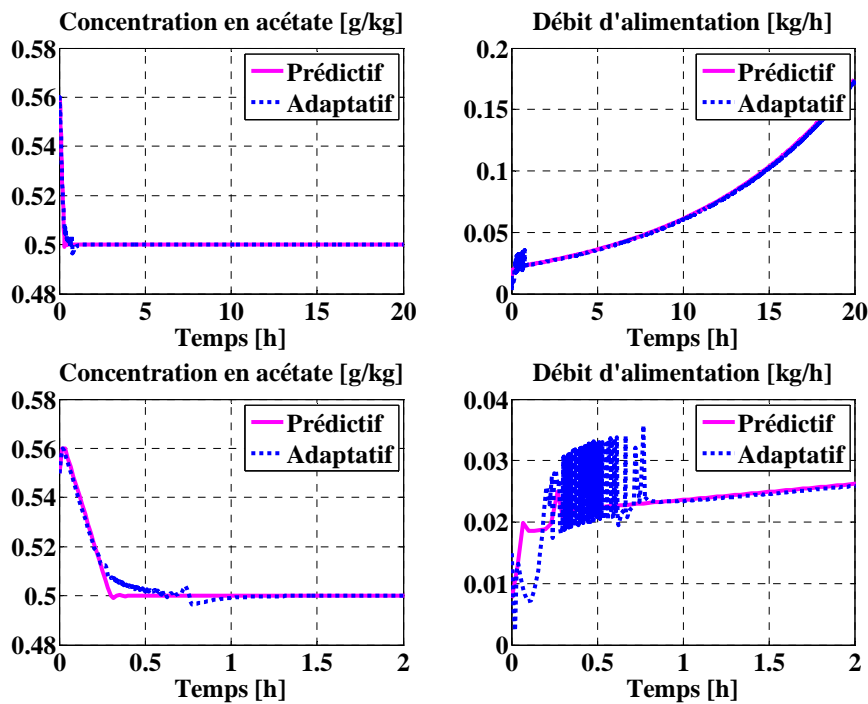
Afin de comparer les performances des deux commandes prédictive et adaptative appliquées à notre bioprocédé, nous présentons ici des résultats de simulation pour la régulation de la

concentration en acétate à la valeur  $A_{set} = 0,5$  g/kg . Les conditions initiales et les paramètres de commande choisis pour ces simulations sont présentés dans le Tableau 5.4.

Types de commande	Commande adaptative linéarisante	Commande prédictive non-linéaire
<b>Conditions initiales</b>	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ O_0 \\ C_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,55 \text{ g/kg} \\ 7,24 \text{ g/kg} \\ 5,59 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,55 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$
<b>Paramètres de commande</b>	Gain de commande : $\gamma = 6$ , soit $Tr_{5\%} = 30$ min	Horizon de prédiction : $N = 8$ Facteur de pondération : $\eta = 0,1$

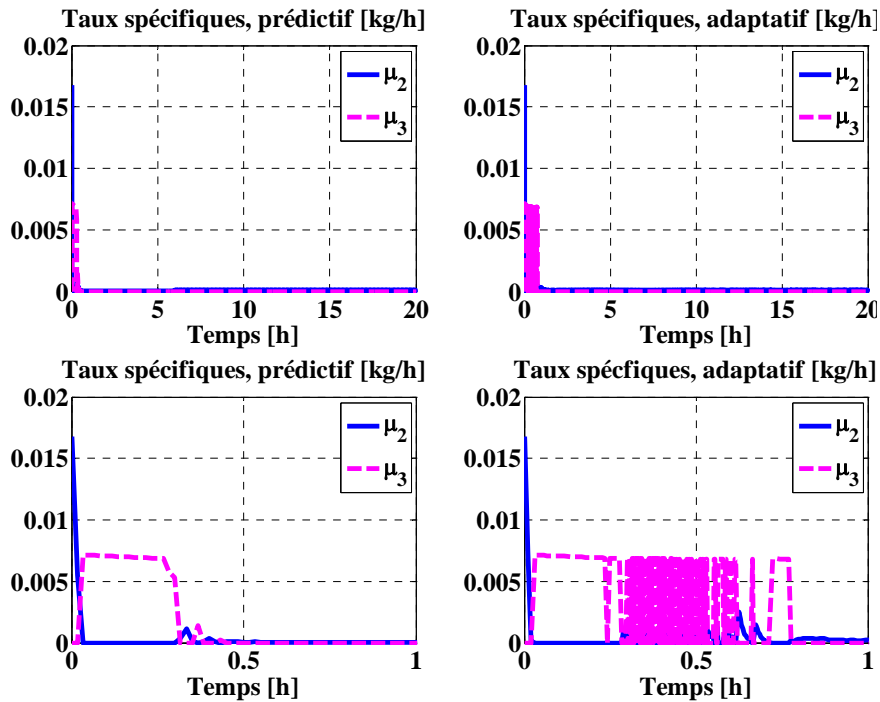
**Tableau 5.4.** Paramètres de réglage des commandes adaptative et prédictive.

La Figure 5.24 présente les évolutions de la concentration en acétate (variable à réguler) et le débit d'alimentation (variable de commande) en utilisant les deux stratégies de commande : prédictive et adaptative.



**Figure 5.24.** Evolution de la concentration en acétate et du débit d'alimentation, commande prédictive et régulateur adaptatif.

Cette figure montre une bonne régulation de la concentration en acétate pour les deux commandes à la valeur  $A_{set} = 0,5$  g/kg. Malgré tout, la commande adaptative linéarisante induit une phase transitoire plus longue et un débit d'alimentation beaucoup plus chahuté que celui issu de la commande prédictive. Pour expliquer ces fluctuations, la Figure 5.25 présente les évolutions des taux spécifiques  $\mu_2$  et  $\mu_3$  pour les deux commandes prédictive et adaptative. Comme il a été indiqué précédemment, les valeurs de ces taux déterminent le régime de fonctionnement du bioprocédé. Si  $\mu_2 = 0$ , on se trouve en régime oxydatif et si  $\mu_3 = 0$ , c'est le régime oxydo-fermentatif qui est actif.

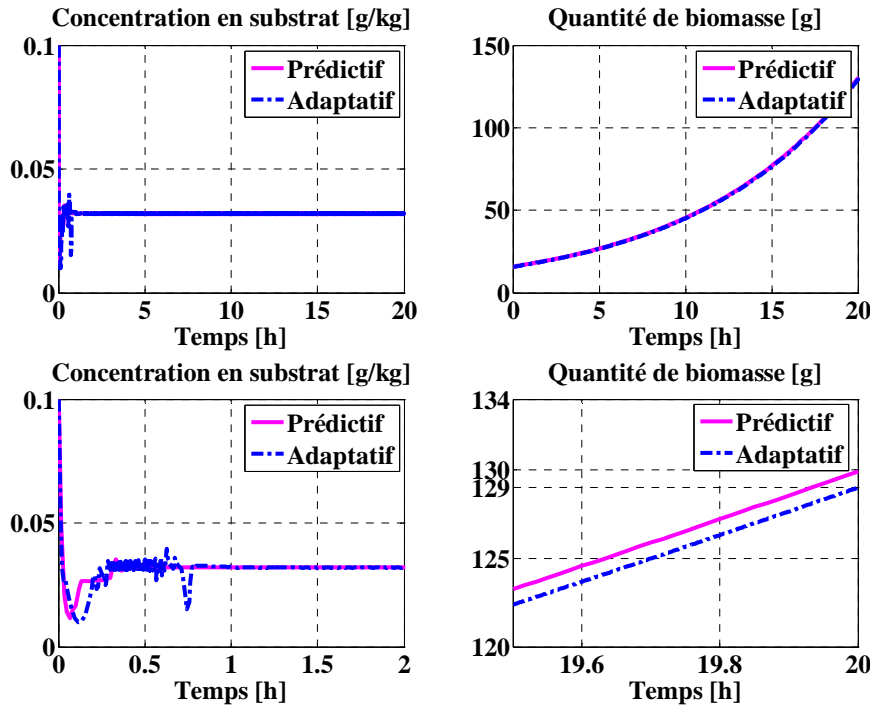


**Figure 5.25.** Evolutions des taux spécifiques  $\mu_2$  et  $\mu_3$ , commande prédictive et régulateur adaptatif.

On constate clairement que la commande adaptative linéarisante induit un fonctionnement du bioprocédé avec de nombreux basculements entre les deux régimes pendant les 40 premières minutes. Ceci se reflète effectivement dans le cas adaptatif sur l'évolution du débit d'alimentation (Figure 5.24).

Rappelons par ailleurs que nous avons mis en évidence la régulation implicite autour d'un couple  $(A, S)$ . Ainsi, la régulation de la concentration en acétate à la valeur  $A_{set} = 0,5$  g/kg implique une régulation de la concentration en substrat à la valeur  $S_{crit} = 0,0319$  g/kg. Les évolutions de la concentration en substrat en utilisant les deux commandes de la Figure 5.26 montrent que la commande adaptative régule également la concentration en substrat à la même valeur  $S_{crit} = 0,0319$  g/kg mais avec une phase transitoire plus longue par rapport à celle issue de l'application de la commande prédictive.

Cette figure présente également les évolutions de la quantité de biomasse pour les deux commandes. L'intérêt ici est de comparer sa valeur à l'instant final pour les deux commandes. On remarque que la quantité de biomasse obtenue à partir de la commande prédictive est légèrement supérieure à celle obtenue avec la commande adaptative.



**Figure 5.26.** Evolution de la concentration en substrat et de la quantité de biomasse, commande prédictive et régulateur adaptatif.

Ainsi, au nominal, la commande adaptative linéarisante développée par Rocha et la commande prédictive non-linéaire élaborée dans le cadre de ces travaux de thèse répondent à l'objectif de commande. Globalement, elles induisent des comportements relativement similaires. Malgré tout, les transitoires issus de la commande adaptative apparaissent beaucoup plus chahutés et plus lents, la croissance de la biomasse est également moins forte. Par ailleurs, la commande prédictive se base sur un modèle d'ordre plus faible (4 variables d'état au lieu de 6), nécessitant un plus faible nombre de paramètres à identifier.

### 5.6.2 Comparaison avec la commande par modèle générique (GMC)

La commande par modèle générique a été développée par Lee et Sullivan à la fin des années 80 [Lee et Sullivan, 1988], et depuis cette commande a été appliquée dans plusieurs travaux, notamment dans le domaine de la biotechnologie. Le principe de cette commande est d'imposer une trajectoire de référence à la variable à réguler. Elle s'écrit sous la forme [Lee et Sullivan, 1988] :

$$\left(\frac{dy}{dt}\right)^* = G_1(y^* - y) + G_2 \int_0^t (y^* - y) dt \quad \text{avec : } G_1 > 0, G_2 > 0 \quad (5.33)$$

où :

- $\left(\frac{dy}{dt}\right)^*$  représente la trajectoire de référence de la dérivée de la variable à réguler,
- $y^*$  représente la consigne,
- $y$  représente la variable à réguler,

- $G_1$  et  $G_2$  sont les paramètres de réglage du régulateur GMC. Ces paramètres sont généralement choisis en se basant sur le fait que la fonction de transfert équivalente en boucle fermée correspond à un transfert du deuxième ordre, d'amortissement  $\xi$  et de pulsation propre  $\omega_0$ , par la relation :

$$\begin{cases} G_1 = 2\xi\omega_0 \\ G_2 = \omega_0^2 \end{cases} \text{ avec : } G_1 > 0, G_2 > 0$$

La commande GMC inclut des actions proportionnelle et intégrale avec une linéarisation entrée/sortie du système, cette commande permet donc d'éliminer l'erreur statique et d'avoir une bonne précision.

### 5.6.2.1 Commande par modèle générique appliquée au bioprocédé *E. coli*

La commande par modèle générique est appliquée au bioprocédé *E. coli* pour la régulation de la concentration en acétate en agissant sur le débit d'alimentation  $F_{in}$  et en utilisant le modèle de commande (5.5).

Reprenons l'équation différentielle de l'évolution de la concentration en acétate (5.15.a) :

$$\frac{dA}{dt} = r_A X - \frac{A}{W} F_{in} \quad (5.34)$$

En imposant à l'évolution de la concentration en acétate la trajectoire de référence (5.33), l'expression du débit d'alimentation est :

$$F_{in} = \left( r_A X - G_1(A_{set} - A) - G_2 \int_0^t (A_{set} - A) dt \right) \frac{W}{A} \quad (5.35)$$

### 5.6.2.2 Résultats de simulation

Nous présentons ici des résultats de simulation permettant la comparaison de notre stratégie de commande à la commande GMC appliquée au bioprocédé. L'objectif fixé est la régulation de la concentration en acétate à la valeur  $A_{set} = 0,5$  g/kg. Les paramètres de réglage choisis pour ces simulations sont présentés dans le Tableau 5.5.

Types de commande	Commande par modèle générique	Commande prédictive non-linéaire
<b>Conditions initiales</b>	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,55 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,55 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$
<b>Paramètres de commande</b>	$G_1 = 12$ , $G_2 = 36$ , ou $\xi = 1$ $\omega_0 = 6$ rad/h, $t_r = 30$ min	Horizon de prédiction : $N = 8$ Facteur de pondération : $\eta = 0,1$

**Tableau 5.5.** Paramètres de réglage des commandes GMC et prédictive.

La Figure 5.27 présente les évolutions de la concentration en acétate et du débit d'alimentation en utilisant les deux commandes GMC et prédictive. Cette figure montre que l'objectif de commande est atteint. Toutefois, le débit d'alimentation issu de la commande par modèle générique a des valeurs beaucoup plus importantes par rapport au débit d'alimentation issu de la commande prédictive.

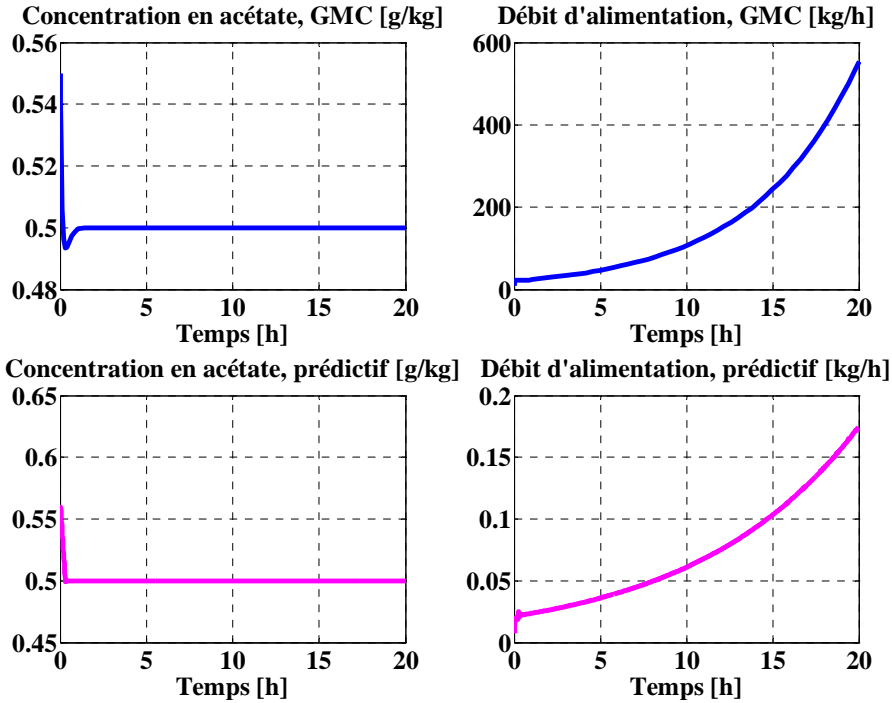


Figure 5.27. Evolution de la concentration en acétate et du débit d'alimentation, commande prédictive et GMC.

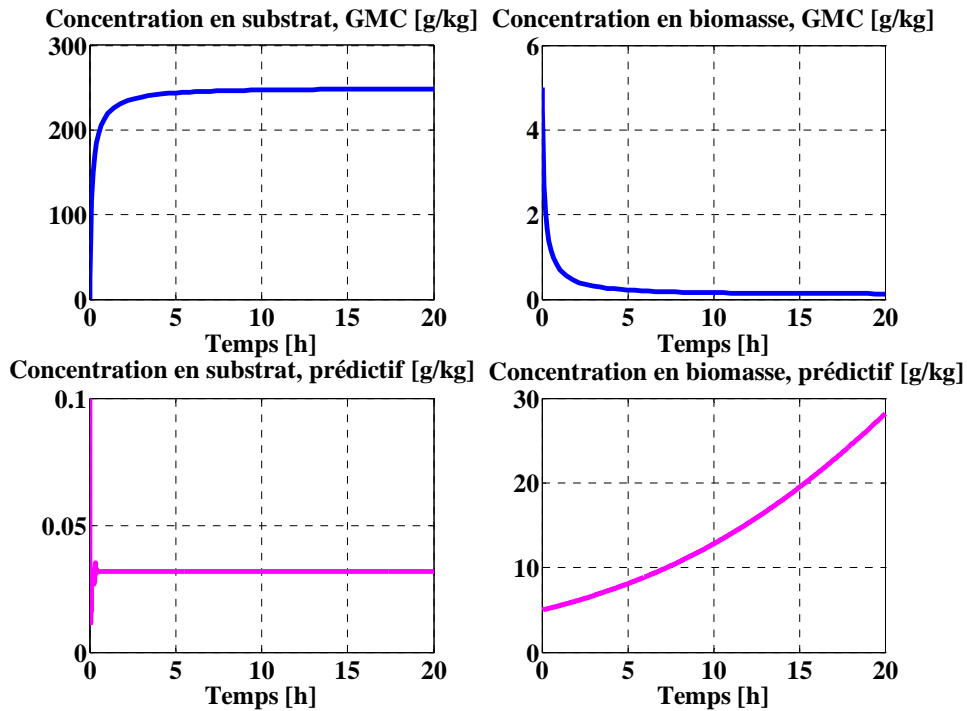


Figure 5.28. Evolution de la concentration en substrat et en biomasse, commande prédictive et GMC.

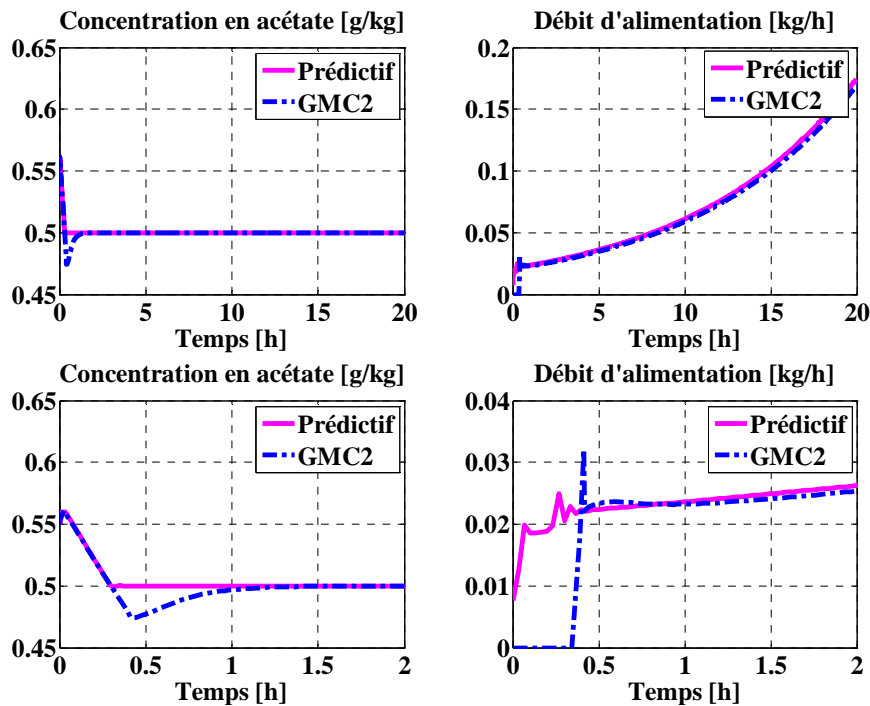
Dès lors que le bioprocédé est alimenté en substrat, des valeurs importantes du débit d'alimentation impliquent alors d'importantes valeurs de concentration en substrat. La Figure 5.28 présente les évolutions des concentrations en substrat et en biomasse. On voit qu'en utilisant la commande par modèle générique, la concentration en substrat est régulée à la valeur 248,34 g/kg, alors qu'en utilisant la commande prédictive, elle est régulée à la valeur  $S_{crit} = 0,0319$  g/kg. Donc, selon la loi de commande utilisée, les concentrations en acétate et en substrat peuvent être régulées au couple  $(A_{set} = 0,5$  g/kg,  $S_{crit} = 0,0319$  g/kg) ou  $(A_{set} = 0,5$  g/kg,  $S_{crit} = 248,34$  g/kg). Mathématiquement, ces deux cas sont possibles (relation (5.17)), mais le deuxième cas n'est pas plausible en pratique. On remarque d'ailleurs que l'évolution de la concentration en biomasse obtenue à l'aide de la GMC n'a pas une forme exponentielle et qu'elle tend vers 0. Ceci n'est pas un profil réaliste, et de surcroît non optimal. On constate donc qu'avec la loi de commande GMC élaborée à partir du modèle simplifié de commande, il est impossible de satisfaire les objectifs de régulation du couple  $(A, S)$ , la solution conduisant à des concentrations et des débits irréalistes.

Envisageons alors d'appliquer la GMC sur le modèle complet, de façon similaire à ce qui a été réalisé dans le cadre de la commande adaptative linéarisante, en partant de nouveau de la relation :

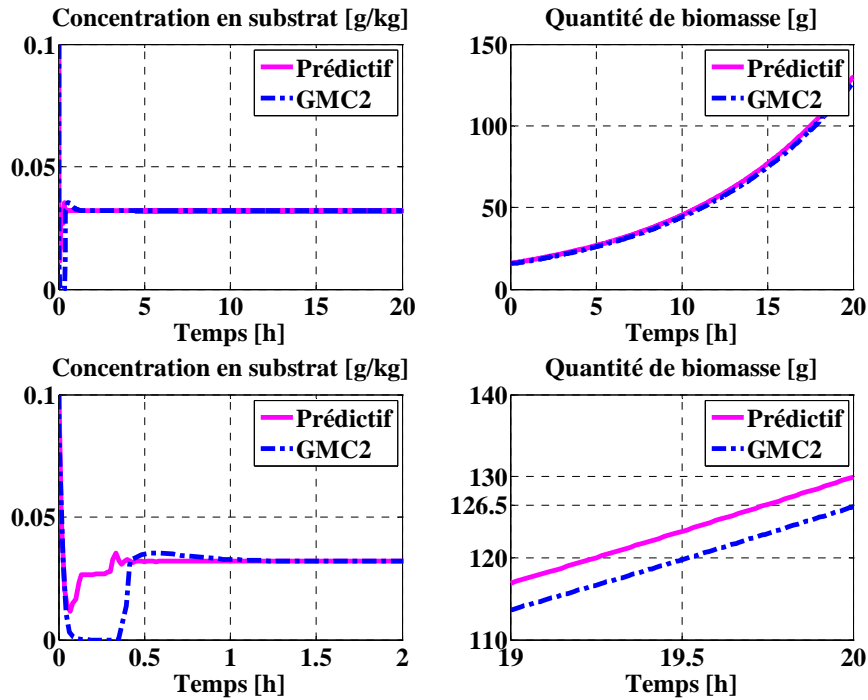
$$\frac{dA}{dt} = -p_1 CTR - p_2 OTR + p_3 D S_{in} - D A \quad (5.36)$$

Dans ce cas, l'expression du débit d'alimentation devient :

$$F_{in} = \frac{p_1 CTR + p_2 OTR + G_1(A_{set} - A) + G_2 \int_0^t (A_{set} - A) dt}{p_3 S_{in} - A} W \quad (5.37)$$



**Figure 5.29.** Evolution de la concentration en acétate et du débit d'alimentation, commande prédictive et GMC (avec modèle complet).



**Figure 5.30.** Evolution de la concentration en substrat et de la quantité de biomasse, commande prédictive et GMC (avec modèle complet).

Les résultats pour le même scénario sont alors donnés par les Figures 5.29 et 5.30. Dans ce cas, le couple  $(A, S)$  recherché est effectivement atteint, mais avec des transitoires beaucoup plus irréguliers, en particulier avec une concentration du substrat passant temporairement par 0 avant de revenir à la valeur souhaitée.

Cette figure présente également les évolutions de la quantité de biomasse pour les deux commandes, prédictive et GMC. Comme dans le cas de la commande adaptative, la quantité de biomasse obtenue à partir de la commande prédictive est supérieure à celle obtenue avec la commande GMC.

En conclusion, pour cette commande comme pour la commande adaptative linéarisante, les performances de la régulation sont moins bonnes, la quantité de biomasse produite pour un même scénario est inférieure et il est nécessaire de se baser sur le modèle le plus complexe alors que la commande prédictive utilise le modèle d'ordre plus faible (4 variables d'état au lieu de 6), nécessitant donc un plus faible nombre de paramètres à identifier.

#### Remarque :

Appliquer avec le modèle complet une commande par modèle générique en posant  $G_1 = \gamma$  et  $G_2 = 0$  revient à considérer la stratégie de commande adaptative linéarisante.

## 5.7 Conclusions

Ce chapitre a proposé la mise en œuvre d'une stratégie de commande prédictive non-linéaire pour la commande du bioréacteur de culture de bactéries *E. coli*. La stratégie considérée au niveau du bioréacteur consiste à maximiser la croissance de la biomasse. Ceci s'est traduit par

la régulation de la concentration en acétate autour d'une valeur aussi proche que possible de zéro compte-tenu de la sensibilité des capteurs, tout en forçant le débit d'alimentation à suivre dans le même temps une commande de référence spécifiée. Pour cela, ce débit d'alimentation de référence a été calculé hors ligne, et dépend de la configuration d'équilibre du couple  $(A, S)$ . En ce sens, l'un des apports de ce chapitre a été de montrer que réguler la concentration en acétate impliquait de façon implicite la régulation en substrat. Un autre apport a été l'élaboration d'une structure prédictive incluant une philosophie de type modèle de référence, puisque le critère prend en compte non seulement des écarts de sortie mais également des écarts de commande.

La synthèse de la commande s'est basée pour la prédiction du comportement futur sur le modèle de commande simplifié défini au chapitre 2. Afin d'éviter la discrétisation de ce modèle, et l'augmentation associée de la charge de calcul temps réel, le problème d'optimisation a été transformé en un problème de programmation non-linéaire par des techniques de CVP, un changement de variable a enfin permis de se ramener à un problème d'optimisation non contraint. La structure élaborée tient compte de plus des erreurs de modèle en incluant lors de la prédiction un signal d'erreur correspondant à la différence entre le système et le modèle.

Ce chapitre a ensuite permis de valider cette stratégie en simulation pour différents scénarii et dans une configuration nominale de paramètres (aucune désadaptation entre le modèle et le bioréacteur réel). Les résultats obtenus montrent de bonnes performances pour la régulation d'acétate, avec des transitoires relativement courts. On constate également que le débit d'alimentation suit le profil exponentiel imposé de façon très satisfaisante. Des indications ont été apportées concernant le choix des paramètres de réglage de la loi de commande. On s'aperçoit que la philosophie à modèle de référence procure une relative insensibilité à ces paramètres, ce qui rend leur choix moins délicat.

Enfin, les performances de la loi de commande élaborée ont été comparées à celles obtenues avec deux stratégies utilisées dans ce domaine, la commande adaptative linéarisante et la commande par modèle générique GMC. Par rapport à la commande prédictive, ces deux méthodes ont montré au nominal, pour des scénarii identiques, de moins bonnes performances, en termes de suivi de trajectoire, de quantité de biomasse produite et de complexité du modèle requis.

Le chapitre suivant va maintenant proposer une étude de la robustesse de la loi de commande prédictive non-linéaire élaborée, en faisant intervenir des désadaptations entre le modèle et le système. On montrera en particulier l'intérêt de la différence objet-modèle et on comparera la robustesse de notre stratégie à celles des deux autres structures de commande mentionnées ci-dessus.

## 6. Etude de la robustesse de la loi de commande prédictive non-linéaire

### 6.1 Introduction

Les chapitres précédents ont permis d'aboutir à la synthèse d'une loi de commande prédictive non-linéaire régulant la concentration en acétate et forçant le débit d'alimentation à suivre un débit de référence spécifié, afin de garantir globalement une maximisation de la croissance de la biomasse. Les performances de la stratégie élaborée ont été illustrées par l'intermédiaire de plusieurs scénarii, et comparées aux résultats fournis par deux autres structures de commande rencontrées dans le cadre de la culture fed-batch des *E. coli*.

La loi de commande proposée a donc jusqu'ici été validée en simulation dans une configuration nominale, en supposant qu'il n'y avait pas de désadaptation entre le système – le bioréacteur réel – et le modèle utilisé pour la synthèse de la commande. Or, dès lors que l'objectif final est l'implantation de la structure de commande sur un bioréacteur réel, de nombreuses sources d'incertitudes et de perturbations vont se manifester, en particulier une mauvaise connaissance des paramètres du système et l'influence de bruits de mesures. En effet, en pratique, le modèle ne reflète jamais parfaitement le système, surtout dans le domaine des bioprocédés. En particulier, comme nous l'avons précisé lors des chapitres précédents, la modélisation et l'identification des bioprocédés en général restent une tâche difficile, et les modèles de commande utilisés sont toujours simplifiés.

La loi de commande élaborée doit donc fournir certaines garanties de robustesse intrinsèque vis-à-vis de ces incertitudes et perturbations. L'objectif de ce chapitre – volontairement séparé du précédent pour insister sur l'importance de la démarche – est de fournir une analyse de la robustesse de la structure de commande, d'une part vis-à-vis d'erreurs de modèle, d'autre part vis-à-vis de l'influence de bruits de mesures.

Pour cela, des résultats de simulation sont tout d'abord présentés afin de comparer les évolutions des grandeurs significatives (concentrations en acétate, glucose, biomasse et débit d'alimentation) pour la configuration nominale puis en désadaptant les valeurs des paramètres. Cette comparaison permettra également d'illustrer l'importance de la différence objet-modèle (D.O.M) intégrée dans la loi de commande du chapitre 5. Dans un deuxième temps, la robustesse vis-à-vis d'erreurs de modèle est abordée par le biais d'une analyse statistique effectuée selon une approche de type Monte Carlo.

L'influence de bruits de mesures est ensuite analysée dans des conditions défavorables, c'est-à-dire en supposant toujours la présence de désadaptation entre le système et le modèle. Tous les résultats mis en évidence doivent permettre de valider la robustesse de la loi de commande.

Une dernière partie se penche sur les performances, en termes de robustesse vis-à-vis d'erreurs de modèle, des deux lois de commande présentées au chapitre 5 (commande adaptative linéarisante et commande par modèle générique) permettant ainsi de les situer par rapport à la stratégie prédictive élaborée dans ce mémoire.

## 6.2 Robustesse vis-à-vis d'erreurs de modèle

Le but de cette partie est de montrer le comportement de la loi de commande prédictive non-linéaire proposée en présence d'une désadaptation importante entre le modèle et le système. Les résultats obtenus doivent permettre d'analyser la robustesse de la structure de commande.

### 6.2.1 Analyse d'un scénario particulier

L'objectif affiché ici, comme pour le chapitre précédent, est de réguler la concentration en acétate à la valeur  $A_{set} = 0,5 \text{ g/kg}$  en imposant également au débit d'alimentation de suivre le profil de référence élaboré au chapitre 5. Les résultats obtenus ont pour but de comparer le comportement de la régulation au nominal et en désadapté.

#### 6.2.1.1 Prise en compte de la différence objet-modèle

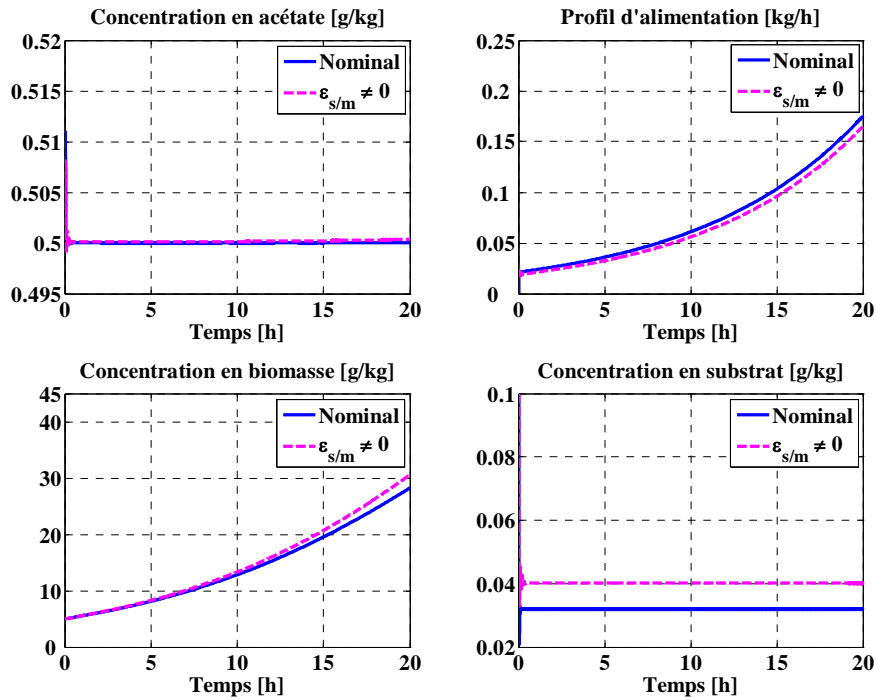
On envisage dans ce paragraphe le scénario et les conditions ci-dessous :

- la période d'échantillonnage est  $T_e = 2 \text{ min}$ ,
- la durée de simulation est égale à 20 heures,
- les paramètres de commande sont :
  - l'horizon de prédiction  $N = 8$ ,
  - le facteur de pondération  $\eta = 0,1$ ,
- les paramètres du modèle ( $k_1, k_2, k_3, k_4, k_s, k_{os}, k_{io}, k_a, k_{ia}, k_{oa}, q_{smax}, q_{acmax}, q_{omax}$ ) sont désadaptés de manière aléatoire entre  $-20\%$  et  $+20\%$ ,
- les valeurs initiales des variables d'état sont choisies très proches du point de régulation pour éliminer les transitoires  $[X_0 \ S_0 \ A_0 \ W_0] = [5 \text{ g/kg} \ 0,1 \text{ g/kg} \ 0,50 \text{ l/g} \ 3,17 \text{ kg}]$ .

La Figure 6.1 représente les évolutions des variables d'état et du débit d'alimentation résultant de la stratégie de commande proposée au chapitre 5. Rappelons que le signal  $\varepsilon_{s/m}$  correspond à la différence entre la sortie du système réel et le modèle (D.O.M), cette variable est non nulle lorsque l'on désadapte les paramètres du modèle.

Cette figure permet de comparer les réponses en présence d'une désadaptation avec celles obtenues au nominal. On constate que les performances de la régulation de la concentration en acétate restent satisfaisantes malgré la désadaptation des paramètres. La dérive observée à la fin de la durée de simulation reste très faible et sera analysée par la suite. Les évolutions des concentrations en biomasse au nominal et en désadapté ont pratiquement la même allure, en revanche on remarque un décalage non négligeable entre les évolutions de la concentration en substrat au nominal et en désadapté. Comme nous l'avons vu précédemment, au nominal, la concentration en substrat est régulée à la valeur  $S_{crit} = 0,0319 \text{ g/kg}$ . Cette valeur correspond à la valeur  $A_{set} = 0,5 \text{ g/kg}$  et rappelons-le, vérifie la condition suivante :

$$\left. \frac{q_S}{(S_{in} - S)} \right|_{S=S_{crit}} = \left. \frac{r_A}{A} \right|_{A=A_{set}, S=S_{crit}} \quad (6.1)$$

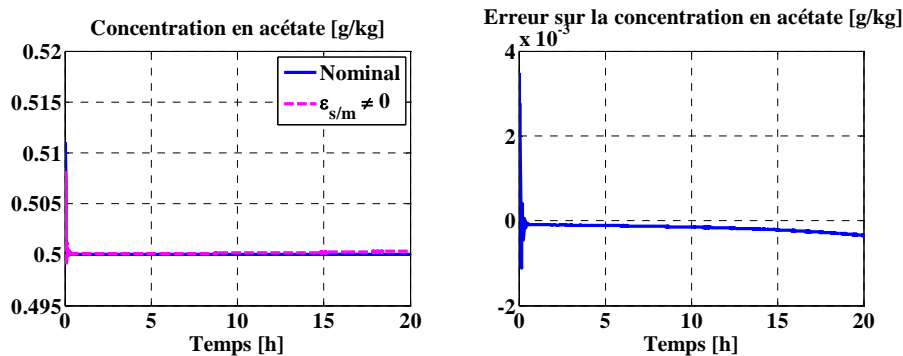


**Figure 6.1.** Evolutions des concentrations en acétate, en substrat et en biomasse et du débit d'alimentation au nominal et en désadapté.

La Figure 6.1 montre que la concentration en substrat issue du modèle désadapté est régulée à une autre valeur de  $S = 0,0401$  g/kg différente de  $S_{crit}$  même si la concentration en acétate est régulée à la valeur  $A_{set} = 0,5$  g/kg. Ce résultat est normal et logique puisque les valeurs des paramètres sont désadaptées, les valeurs des taux  $r_A$  et  $q_S$  ne sont plus les mêmes. Il a été vérifié que la valeur  $S = 0,0401$  g/kg satisfait la relation (6.1) pour  $A_{set} = 0,5$  g/kg et les valeurs désadaptées des paramètres.

Enfin, le débit d'alimentation du cas désadapté suit une allure exponentielle, mais légèrement différente de celle observée au nominal, ce qui s'explique ici encore par le fait que le couple  $(A_{set}, S_{crit})$  ne soit plus exactement celui attendu. Globalement, tous ces résultats montrent que la stratégie de commande proposée présente à ce stade de l'analyse de bonnes propriétés de robustesse et qu'elle arrive à prendre en compte les variations de paramètres.

Pour illustrer plus précisément la différence entre l'évolution de la concentration en acétate au nominal et en désadapté, la Figure 6.2 présente les évolutions de cette concentration dans les deux cas ainsi que l'évolution de la différence des deux courbes (nominal – désadapté).



**Figure 6.2.** Evolution de la concentration en acétate au nominal et en désadapté et évolution de l'erreur entre ces deux réponses.

Cette figure montre que l'erreur sur la concentration en acétate est très faible, de l'ordre de  $0,510^{-4}$  g/kg, sachant que les paramètres du modèle ont été désadaptés de  $\pm 20\%$ . La loi de commande proposée au chapitre 5 permet alors d'atteindre l'objectif fixé qui est de réguler la concentration en acétate malgré une désadaptation des valeurs des paramètres assez importante ( $\pm 20\%$ ).

**Remarque :**

Pour aller plus loin et diminuer encore cette erreur résiduelle sur la concentration en acétate, on peut envisager de répartir la différence objet-modèle sur tout l'intervalle de prédiction selon la relation :

$$\hat{A}_{k+j} = \hat{A}_{\text{mod}_{k+j}} + j \underbrace{(A_k - A_{\text{mod}_k})}_{\varepsilon_{s/m}(k)}, \quad j = \overline{1, N} \quad (6.2)$$

au lieu de la structure utilisée dans la loi de commande proposée :

$$\hat{A}_{k+j} = \hat{A}_{\text{mod}_{k+j}} + \underbrace{(A_k - A_{\text{mod}_k})}_{\varepsilon_{s/m}(k)}, \quad j = \overline{1, N} \quad (6.3)$$

### 6.2.1.2 Utilité de la différence objet-modèle

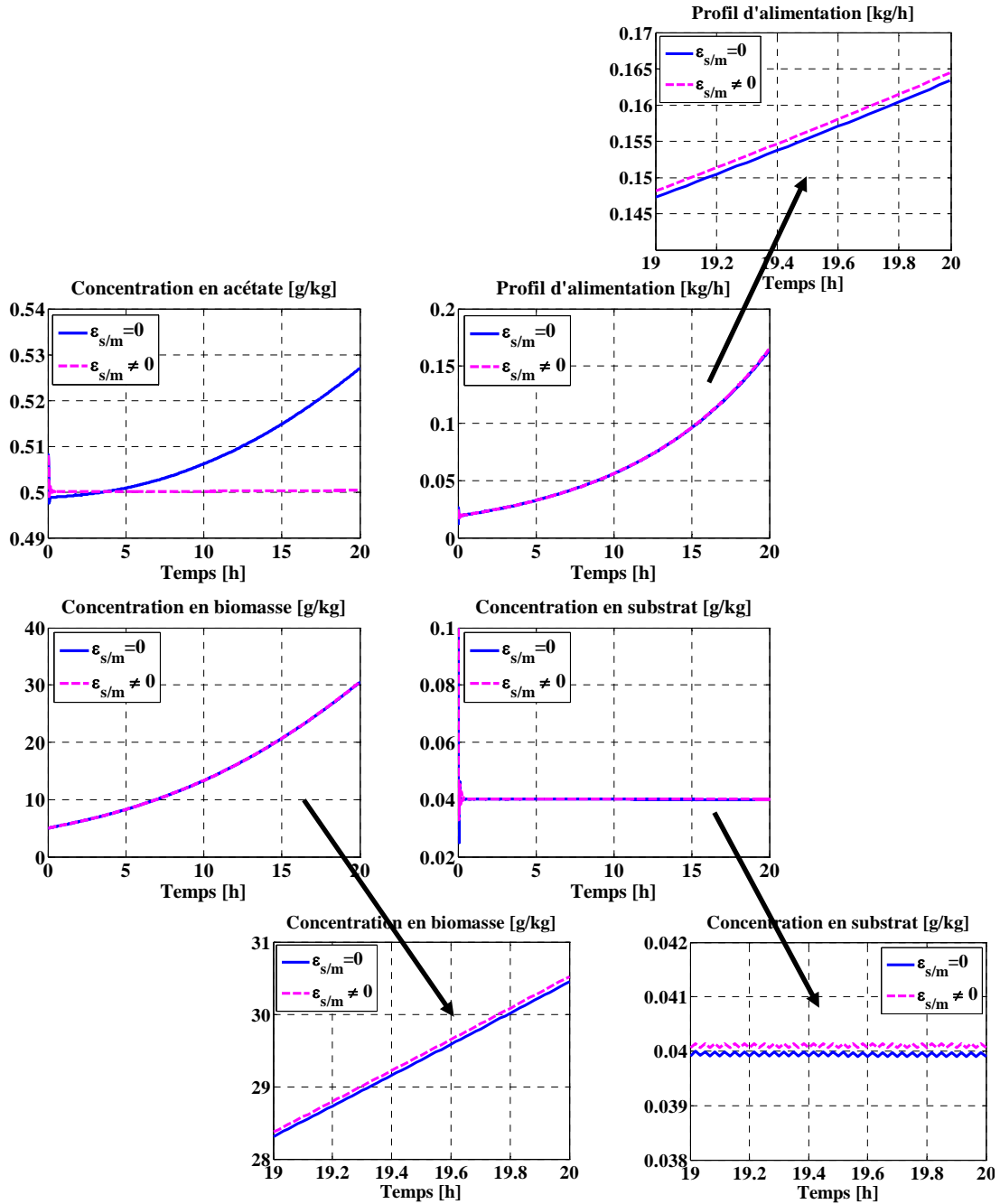
L'objectif de cette partie est de montrer, en analysant les résultats de simulation, l'utilité de la différence objet modèle (D.O.M). Cette différence est prise en compte dans la loi de commande que nous avons proposée pour tenir compte de la différence possible entre le système et le modèle de commande. Au nominal, cette différence est nulle car on considère que le modèle est identique au système. En désadaptant les paramètres du modèle, on quitte la configuration nominale de sorte que, dans ce cas, la variable  $\varepsilon_{s/m}$  n'est plus nulle.

Pour illustrer l'avantage de prendre en compte la différence objet-modèle dans la loi de commande proposée, nous avons réalisé des simulations en partant des mêmes conditions opératoires que précédemment, avec la même désadaptation, pour atteindre le même objectif qui est de réguler la concentration en acétate à la valeur  $A_{\text{set}} = 0,5$  g/kg. Ces simulations sont effectuées dans les cas suivants :

1. Réponses à la loi de commande prédictive non-linéaire sans prendre en compte la différence objet-modèle ( $\varepsilon_{s/m} = 0$ ),
2. Réponses à la loi de commande prédictive non-linéaire en prenant en compte la différence objet-modèle ( $\varepsilon_{s/m} \neq 0$ ). C'est la loi de commande proposée au chapitre 5.

Le scénario choisi ici reprend les mêmes conditions que celles utilisées au paragraphe 6.2.1.1. La Figure 6.3 représente les évolutions des concentrations en acétate, en biomasse et en substrat et le débit d'alimentation dans les deux cas.

Cette figure montre que pour le premier cas ( $\varepsilon_{s/m} = 0$ ), la concentration en acétate diverge ; elle n'est plus régulée, l'objectif de la loi de commande n'est donc pas atteint. Tandis que pour le deuxième cas ( $\varepsilon_{s/m} \neq 0$ ), la concentration en acétate est régulée à la valeur désirée  $A_{\text{set}} = 0,5$  g/kg. Ce résultat est cohérent puisque le fait de considérer la différence objet-modèle non nulle revient à considérer une sorte d'action intégrale implicite qui réduit alors l'erreur entre la sortie et la consigne.



**Figure 6.3.** Evolutions des concentrations en acétate, en biomasse et en substrat et du débit d'alimentation issus de la commande prédictive non-linéaire avec et sans la D.O.M.

Les évolutions des autres variables d'état et du débit d'alimentation sont pratiquement identiques pour les deux cas. Les zooms de ces tracés sont présentés à chaque fois pour distinguer les deux configurations. On constate une légère différence entre les réponses des deux cas pour les concentrations en biomasse et en substrat et pour le débit d'alimentation. La régulation ne fonctionne plus correctement car la prédiction s'effectue sans avoir connaissance de la désadaptation, et ne permet pas de se recalibrer sur un profil d'alimentation qui correspondrait au couple  $(A_{set}, S_{crit})$  adéquat.

Le tracé de la concentration en biomasse pour  $\varepsilon_{s/m} = 0$  est au-dessous du tracé pour  $\varepsilon_{s/m} \neq 0$ , ceci traduit l'effet inhibiteur de l'acétate sur l'évolution de la biomasse. En effet,

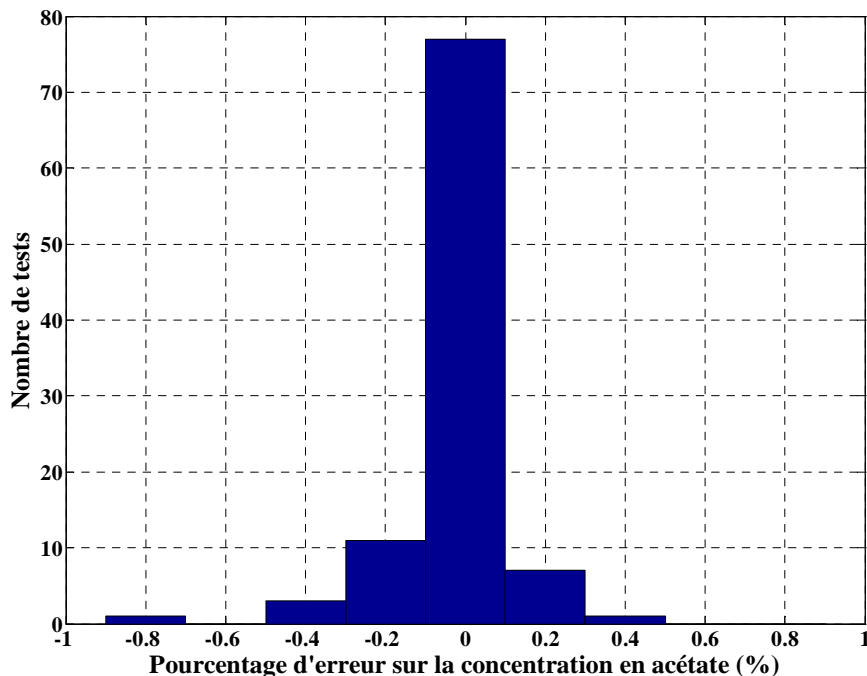
dans le premier cas, la concentration en acétate diverge et est supérieure à la valeur  $A_{set} = 0,5 \text{ g/kg}$ .

## 6.2.2 Analyse statistique

Nous avons vu ci-dessus sur un scénario donné que l'introduction de la différence objet-modèle dans la loi de commande permet de garantir les performances de la régulation de la concentration en acétate en termes de robustesse vis-à-vis d'erreurs paramétriques. Afin de souligner cet aspect, une analyse statistique de robustesse est maintenant envisagée, selon une approche numérique aléatoire de type Monte Carlo, du nom de la roulette de Monte Carlo [Hammersley et Handscomb, 1967].

Pour appliquer cette démarche, 100 tests ont été réalisés en simulation en désadaptant les paramètres de manière aléatoire et non corrélée entre  $-20\%$  et  $+20\%$ . Les conditions opératoires choisies sont identiques à celles présentées au paragraphe 6.2.1. Les résultats de ces tests sont illustrés par l'histogramme présentant la répartition des tests selon la valeur du pourcentage d'erreur relative sur la concentration en acétate :  $(A - A_{set})/A_{set}$ .

La Figure 6.4 représente cet histogramme. On constate que l'on obtient un pourcentage d'erreur compris entre  $-0,1\%$  et  $0,1\%$  pour 77% des tests réalisés. Ceci signifie que pour 77 tests, l'erreur  $A - A_{set}$  est comprise entre  $-5 \cdot 10^{-4} \text{ g/kg}$  et  $5 \cdot 10^{-4} \text{ g/kg}$ , ce qui est suffisamment faible pour donner une vision favorable de la robustesse de notre loi de commande. Notons que la valeur moyenne du pourcentage d'erreur relative obtenue à partir de cet histogramme est de  $-0,0383\%$  et l'écart type est de  $0,129\%$ .



**Figure 6.4.** Histogramme du pourcentage d'erreur relative sur la concentration en acétate pour des variations aléatoires des paramètres de 20%.

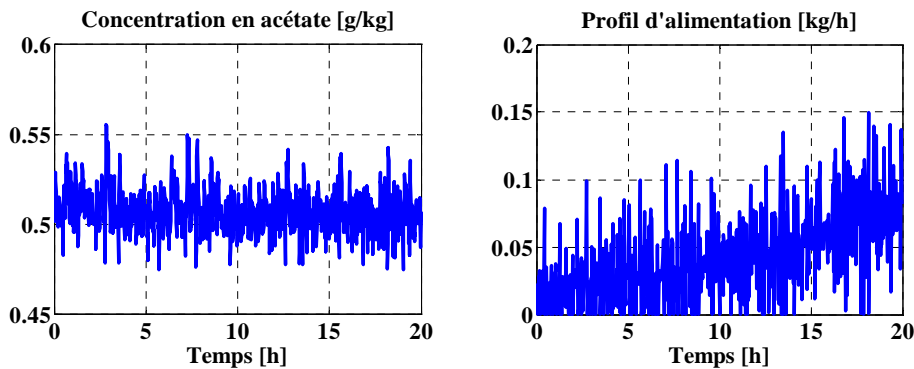
Cette méthode nous permet ainsi de valider la bonne robustesse de la loi de commande proposée au chapitre 5 à l'aide d'outils statistiques.

### 6.3 Analyse de la robustesse vis-à-vis du bruit de mesure

Après avoir étudié la robustesse de notre loi de commande vis-à-vis des variations de paramètres, nous nous intéressons maintenant à l'aspect robustesse vis-à-vis des bruits de mesures issus des différents capteurs. Pour cela, le scénario de simulation suivant est mis en œuvre en simulation :

- l'objectif est de réguler la concentration en acétate à la valeur  $A_{set} = 0,5 \text{ g/kg}$ ,
- la période d'échantillonnage est  $T_e = 2 \text{ min}$ ,
- la durée de simulation est de 20 heures,
- les paramètres de commande sont :
  - l'horizon de prédiction  $N = 8$ ,
  - le facteur de pondération  $\eta = 0,1$ ,
- les paramètres du modèle ( $k_1, k_2, k_3, k_4, k_s, k_{os}, k_{io}, k_a, k_{ia}, k_{oa}, q_{smax}, q_{acmax}, q_{omax}$ ) sont désadaptés de manière aléatoire entre  $-20\%$  et  $20\%$ ,
- les mesures des quatre variables d'état ( $X, S, A$  et  $W$ ) sont bruitées avec un bruit de moyenne nulle et d'écart type égal à  $0,01$ ,
- les valeurs initiales des variables d'état du modèle sont :
 
$$[X_0 \ S_0 \ A_0 \ W_0] = [5 \text{ g/kg} \ 0,1 \text{ g/kg} \ 0,501 \text{ g/kg} \ 3,17 \text{ kg}].$$

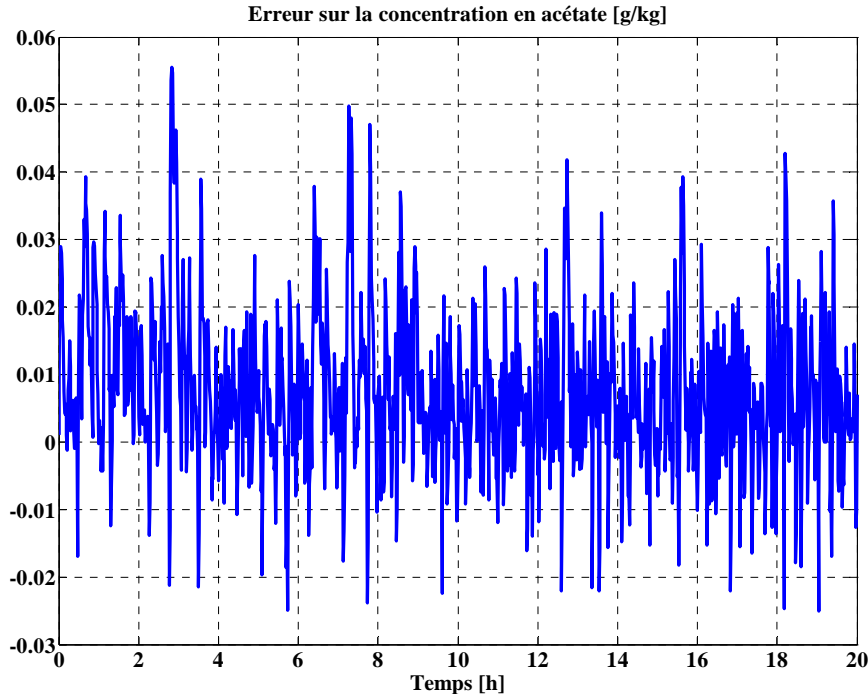
La Figure 6.5 présente les évolutions de la concentration en acétate et du débit d'alimentation pour le scénario décrit ci-dessus.



**Figure 6.5.** Evolution de la concentration en acétate et du débit d'alimentation.

La concentration en acétate obtenue est bruitée et varie autour de la valeur  $0,5075 \text{ g/kg}$ . Cette valeur est proche de la valeur désirée  $A_{set} = 0,5 \text{ g/kg}$ . Cette différence vient du fait que le modèle est désadapté, le choix de la désadaptation se faisant toujours de manière aléatoire. Le débit d'alimentation est également très bruité à cause de la prise en compte des bruits de capteurs. On retrouve de plus ici l'un des inconvénients de la commande prédictive, qui a tendance à amplifier les bruits haute fréquence (typiquement les bruits de mesures) et induire une répercussion forte sur la commande. Ce phénomène peut être fortement atténué par des techniques de robustification de la loi de commande vis-à-vis des bruits (filtrage des mesures avec prise en compte de ce filtre dans la prédiction).

La Figure 6.6 présente l'erreur sur la concentration en acétate  $A - A_{set}$ , la moyenne de cette erreur est de 0,0075 g/kg (erreur relative de 1,5%), résultat moins bon que lors de l'étude statistique précédente mais malgré tout acceptable, et son écart type est de 0,0118, valeur qui se rapproche du bruit de mesure utilisé pour la simulation. Comme mentionné ci-dessus, la moyenne est non nulle car les paramètres du modèle sont ici désadaptés.



**Figure 6.6.** Evolution de l'erreur sur la concentration en acétate.

Ainsi, malgré les bruits de mesures issus des capteurs et la variation des valeurs des paramètres, la loi de commande proposée permet de converger avec une très bonne précision vers la valeur de la concentration en acétate désirée.

## 6.4 Comparaison avec d'autres commandes

Nous avons montré à l'aide d'une analyse statistique la robustesse de notre loi de commande développée au chapitre 5. Nous allons nous intéresser pour terminer ce chapitre à la comparaison de la robustesse vis-à-vis des paramètres de la loi prédictive par rapport à ce que donneraient les deux stratégies de commande présentées au chapitre 5 (commande adaptative linéarisante [Rocha, 2003] et commande par modèle générique [Lee et Sullivan, 1988]).

### 6.4.1 Comparaison avec la commande adaptative linéarisante

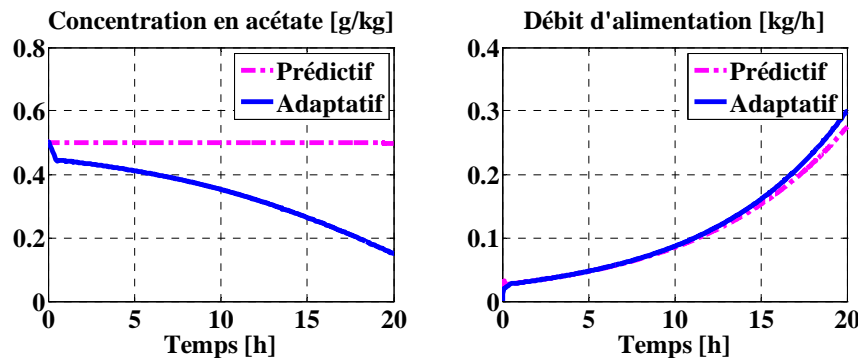
L'application de cette commande au bioprocédé *E. coli* a été présentée au paragraphe 5.6.1.1 du chapitre 5. Au nominal, nous avons pu comparer les performances de cette loi de commande à celle que nous proposons, pour s'apercevoir que la loi prédictive donnait des performances globalement plus satisfaisantes. On s'intéresse ici à comparer les performances en robustesse vis-à-vis des variations de paramètres. Pour cela, deux exemples de simulation

sont présentés, l'un pour l'application de la loi de commande prédictive et l'autre pour l'application de la loi de commande adaptative linéarisante. L'objectif fixé pour ces scénarii est de réguler la concentration en acétate à la valeur  $A_{set} = 0,5 \text{ g/kg}$ . Les valeurs des paramètres du modèle sont les mêmes dans les deux cas (prédictif et adaptatif). Ils sont issus d'une désadaptation aléatoire entre -20% et +20%. Les conditions initiales et les paramètres de commande choisis pour ces simulations sont présentés dans le Tableau 6.1.

Types de commande	Commande adaptative linéarisante	Commande prédictive non-linéaire
Conditions initiales	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ O_0 \\ C_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,501 \text{ g/kg} \\ 7,24 \text{ g/kg} \\ 5,59 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,501 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$
Paramètres de commande	Gain de commande : $\gamma = 6$ , soit $Tr_{5\%} = 30 \text{ min}$	Horizon de prédiction : $N = 8$ Facteur de pondération : $\eta = 0,1$

**Tableau 6.1.** Paramètres de réglage des commandes adaptative et prédictive.

La Figure 6.7 représente les évolutions des concentrations en acétate et des débits d'alimentation en utilisant les deux stratégies de commande suivantes : commande prédictive non-linéaire et commande adaptative linéarisante.



**Figure 6.7.** Evolution de la concentration en acétate et du débit d'alimentation, commande prédictive et régulateur adaptatif.

On constate que la concentration en acétate diverge très rapidement dans le cas du régulateur adaptatif, alors qu'elle est régulée dans le cas de notre loi de commande. La loi de commande adaptative linéarisante ne permet pas de réguler la concentration en acétate dans le cas d'un modèle désadapté car elle ne fait pas intervenir d'action intégrale. Le débit d'alimentation issu de la commande adaptative possède à peu près la même forme dans les deux cas. Le résultat obtenu ici à partir de la commande adaptative linéarisante rappelle celui obtenu dans le cas de

la commande prédictive non-linéaire sans la D.O.M ( $\varepsilon_{s/m}$ ) au paragraphe 6.2.1.2 de ce chapitre. Par ailleurs, comme mentionné au chapitre 5, la commande prédictive non-linéaire est synthétisée à partir d'un modèle simplifié à quatre variables d'état alors que la commande adaptative linéarisante nécessite l'emploi du modèle à six variables d'état, et est donc de fait plus sensible à la phase d'identification, puisqu'il faut identifier plus de paramètres.

## 6.4.2 Comparaison avec la commande par modèle générique (GMC)

L'application de la commande par modèle générique au bioprocédé *E. coli* a été présentée au paragraphe 5.6.2.2 tout d'abord en utilisant le modèle de commande à quatre variables d'état, nous avons montré que la loi de commande obtenue ne permet pas dans ce cas de répondre aux objectifs désirés. Nous avons alors envisagé d'appliquer cette stratégie de commande sur le modèle de commande complet à six variables d'état, les résultats obtenus au chapitre 5 sont similaires en terme de régulation à ceux obtenus à partir de la loi de commande proposée dans cette thèse, mais avec des transitoires plus longs et plus irréguliers. Nous allons maintenant nous intéresser aux performances de cette loi de commande en termes de robustesse. Pour cela, nous comparons les résultats de simulation issus des deux commandes : GMC et prédictive. Deux scénarii sont présentés avec les mêmes variations de paramètres. Les conditions initiales et les paramètres de réglage sont présentés dans le Tableau 6.2.

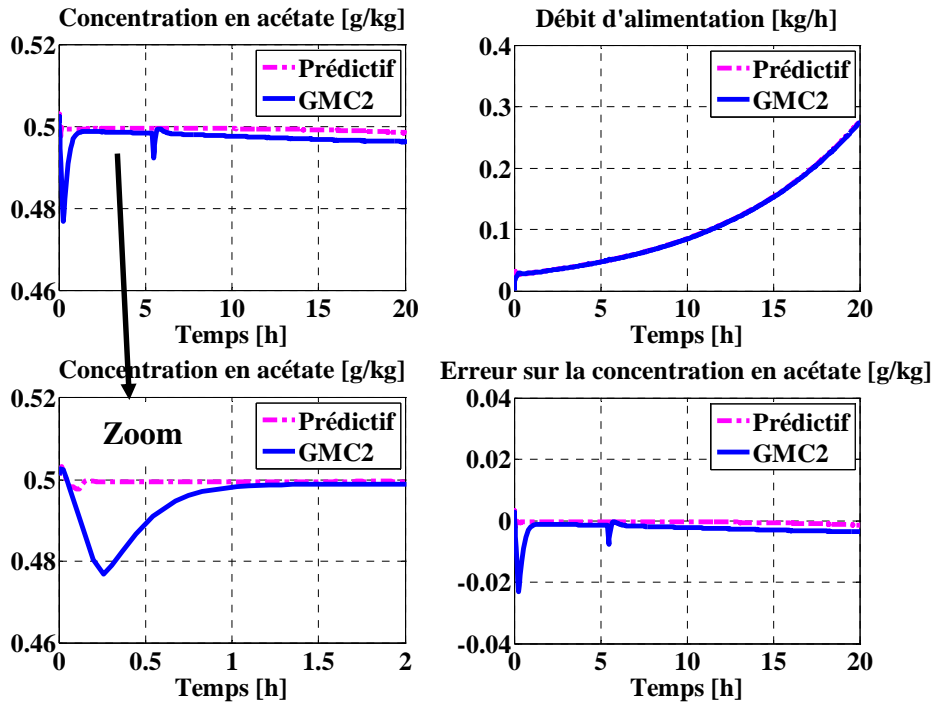
Types de commande	Commande par modèle générique	Commande prédictive non-linéaire
Conditions initiales	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,501 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$	$\begin{bmatrix} X_0 \\ S_0 \\ A_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} 5 \text{ g/kg} \\ 0,1 \text{ g/kg} \\ 0,501 \text{ g/kg} \\ 3,17 \text{ kg} \end{bmatrix}$
Paramètres de commande	$G_1 = 12$ , $G_2 = 36$ , ou $\xi = 1$ $\omega_0 = 6 \text{ rad/h}$ , $t_r = 30 \text{ min}$	Horizon de prédiction : $N = 8$ Facteur de pondération : $\eta = 0,1$

**Tableau 6.2.** Paramètres de réglage des commandes GMC et prédictive.

La Figure 6.8 présente les évolutions des concentrations en acétate et des débits d'alimentation en utilisant les deux stratégies de commande suivantes : commande prédictive et commande par modèle générique (GMC).

Cette figure montre que la concentration en acétate s'écarte de la référence de façon plus importante dans le cas de la GMC, malgré la prise en compte de l'action intégrale dans cette loi de commande. Le zoom montre que, dans ce même cas, le dépassement de la concentration en acétate est très important et le temps de réponse est plus important comparé au comportement de la concentration en acétate issu de la commande prédictive non-linéaire. Cette figure montre également les erreurs sur la concentration en acétate ( $A - A_{set}$ ) dans les deux cas : prédictif et GMC, on constate que l'erreur dans le cas de la GMC est plus grande

que celle obtenue dans le cas du prédictif. Les débits d'alimentation obtenus à partir des deux stratégies de commande sont pratiquement identiques.



**Figure 6.8.** Evolution de la concentration en acétate et du débit d'alimentation, commande prédictive et GMC.

## 6.5 Conclusions

Ce chapitre a présenté une étude de robustesse de la loi de commande prédictive non-linéaire proposée au chapitre 5. L'étude mise en place a fait intervenir des tests en simulation, avec erreurs paramétriques au niveau du modèle et présence de bruits de mesures, complétée par une validation statistique.

L'analyse d'un scénario désadapté a permis de montrer que la loi de commande prédictive garantissait l'objectif de régulation malgré une désadaptation importante des paramètres du modèle. Pour aboutir à ce résultat, la prise en compte de la différence objet-modèle s'est avérée d'un intérêt crucial. Ces conclusions ont ensuite été validées par une étude statistique se basant sur une approche type Monte Carlo. L'étude de robustesse a été également élargie à l'analyse de la robustesse vis-à-vis des bruits de mesures des capteurs, illustrant de même les performances de la loi de commande prédictive non-linéaire à la fois vis-à-vis des bruits de mesures et d'erreurs paramétriques. Enfin, les performances en robustesse de notre loi de commande ont été comparées à celles obtenues avec les deux stratégies présentées au chapitre 5 (commande adaptative linéarisante et commande par modèle générique), illustrant pour des scénarii identiques la supériorité de la loi prédictive.

La commande prédictive non-linéaire élaborée au chapitre 5 présente ainsi un degré de robustesse intrinsèque satisfaisant vis-à-vis de variations paramétriques et de bruits de mesures et pourrait donc être implantée pour piloter un bioréacteur réel avec de bonnes

chances de réussite. Cette implantation, prévue au laboratoire de biochimie de la Faculté Polytechnique de Mons, n'a malheureusement pas pu être réalisée dans le cadre de ces travaux à cause de plusieurs problèmes techniques au niveau de l'installation. Le chapitre suivant de conclusions et perspectives reviendra sur cet aspect.

Pour conclure sur cet aspect de robustesse, remarquons que toute la démarche mise en œuvre ici a consisté à valider a posteriori la robustesse intrinsèque de la loi de commande prédictive non-linéaire, sachant que celle-ci n'est pas à la base une loi de commande robuste à proprement parler. L'analyse menée a permis de conclure à des performances intéressantes en présence d'erreurs paramétriques et de bruits de mesures importants, ce qui constitue déjà un résultat intéressant. Une des perspectives de ce travail serait alors de considérer dès la phase de synthèse des contraintes de robustesse, en termes d'incertitudes, de façon à résoudre un problème d'optimisation de type min-max. Cette méthode est par exemple développée dans [Diehl *et al.*, 2008] et est appliquée dans le cas des colonnes à distiller. Le principe de cette méthode consiste à prendre en compte des incertitudes sur les paramètres du modèle dont on impose les bornes et à choisir le cas le plus défavorable qui sera ensuite minimisé pour la synthèse de la loi de commande. Dans ce cas, on élabore par structure une loi de commande robuste qui peut s'avérer relativement conservative. Il serait alors nécessaire de comparer les performances résultant de la loi développée dans ce mémoire à celles issues de cette structure robuste, en termes de régulation et de robustesse.

## 7. Conclusions - Perspectives

### 7.1 Bilan du travail effectué

Ce travail de thèse a permis de développer, dans le cadre de la culture fed-batch de bactéries *E. coli*, une méthodologie de synthèse d'une loi de commande prédictive non-linéaire, l'objectif principal étant la maximisation de la croissance de la bactérie. Ce travail, mettant en œuvre une démarche générique (modélisation, identification, recherche d'un profil optimal d'alimentation et synthèse de la commande) applicable à d'autres bioprocédés, constitue de fait une étude préliminaire en vue d'une implantation de la loi de commande prédictive non-linéaire à un bioréacteur à l'échelle du laboratoire ou industrielle. La stratégie de commande considérée, ainsi que l'analyse de robustesse proposée, montrent que cette implantation doit pouvoir s'effectuer dans des conditions très favorables. Plusieurs étapes successives, toutes indispensables, ont été examinées :

1. Détermination d'un modèle mathématique macroscopique du bioprocédé, en l'occurrence le bioprocédé de culture des bactéries *E. coli* : Ce modèle est déterminé à partir du schéma réactionnel décrivant les différents métabolismes des bactéries *E. coli*. Etant donné la complexité du modèle déterminé et en vue du développement d'une stratégie de commande avancée, un modèle mathématique simplifié, noté « modèle de commande », est ensuite mis en œuvre, et utilisé pour toutes les étapes aboutissant à la caractérisation de la loi de commande.
2. Identification paramétrique du modèle. Deux stratégies d'identification ont été présentées et validées en simulation, puis comparées à une stratégie d'identification simultanée de l'ensemble des paramètres du modèle. La stratégie d'identification réalisant le meilleur compromis entre simplicité de mise en œuvre et performances se base sur l'analyse de sensibilité des sorties du système vis-à-vis des paramètres du modèle. Elle se déroule en plusieurs étapes, avec à chaque fois les scénarii les plus adaptés à l'ensemble de paramètres recherchés.
3. Détermination d'un profil optimal d'alimentation pour la maximisation de la croissance de la biomasse : Le problème d'optimisation à résoudre s'avère complexe du fait de la nature hybride du modèle du système (existence de deux modes de fonctionnement). Cette difficulté a été contournée en caractérisant toutes les solutions optimales possibles par application du Principe du Maximum de Pontryagin. Enfin, le choix du profil optimal permettant la maximisation de la biomasse a été réalisé à l'aide de résultats numériques et étoffé par une analyse biologique du fonctionnement du bioprocédé. Ainsi, la solution optimale consiste à maintenir le bioréacteur à la frontière des deux régimes – oxydatif et oxydo-fermentatif – maintenant alors la concentration en acétate à zéro. Cette étude a fourni un résultat important utilisé par la suite : la maximisation de la croissance de la biomasse s'obtient en régulant la concentration en acétate à une valeur nulle et en forçant le débit d'alimentation à suivre le profil de référence optimal associé à cette concentration en acétate.
4. Synthèse et application de la commande prédictive non-linéaire au bioprocédé *E. coli* : Comme mentionné ci-dessus, l'objectif est de réguler la concentration en acétate à une valeur nulle en forçant le débit d'alimentation à suivre un profil de

référence, de façon similaire à une commande prédictive à modèle de référence. Notons que cette démarche peut aussi s'apparenter à une approche par la théorie de la platitude, pour laquelle on viendrait superposer une commande en boucle ouverte (le profil d'alimentation optimal par exemple) et un terme complémentaire en boucle fermée issu de la minimisation du critère. Les tentatives de recherche d'une sortie plate physiquement représentative n'ont malheureusement pas abouti dans le cas des bactéries *E. coli*, la complexité du modèle ne facilitant pas les calculs analytiques.

Le problème d'optimisation sous contraintes issu du critère prédictif est ensuite formulé sous la forme d'un problème de programmation non-linéaire sans contraintes, résolu à l'aide d'une approche séquentielle de type CVP, permettant la discrétisation du vecteur de commande uniquement. Pour une meilleure robustesse de la structure proposée, la différence entre le système et le modèle est explicitement incluse dans l'algorithme résultant.

Compte tenu des sensibilités des capteurs, une régulation de la concentration en acétate à une valeur nulle s'avère en fait peu réaliste. Dès lors, la régulation est également envisagée pour une valeur de référence de concentration en acétate non nulle, mais proche de zéro, en faisant intervenir le profil de référence correspondant, calculé pour le couple  $(A_{set}, S_{crit})$  donné. Plusieurs scénarii sont présentés permettant la validation de la stratégie proposée en simulation. Les performances de cette loi de commande sont comparées au nominal à deux autres lois de commandes rencontrées pour ce type de bioprocédé : la commande adaptative linéarisante et la commande par modèle générique. La loi de commande prédictive proposée s'avère globalement plus performante, à la fois en suivi de consigne (même si ce n'est pas la but principal recherché) et en régulation, nécessitant un modèle moins complexe et donc moins de paramètres à identifier.

5. Etude de robustesse : Cette dernière partie – très importante pour juger de l'applicabilité au système réel – se base principalement sur des résultats de simulation et sur des essais statistiques. Le but est de montrer la robustesse de la loi de commande proposée vis-à-vis d'erreurs paramétriques et vis-à-vis de bruits de mesure, reproduisant ainsi la réalité de l'implantation sur un vrai bioréacteur. Les performances de la loi de commande prédictive s'avèrent ici encore meilleures que celles obtenues en utilisant la commande adaptative linéarisante ou la commande par modèle générique.

Les différentes étapes présentées dans ce mémoire ont permis de développer la loi de commande prédictive non-linéaire pour la commande de bioprocédé de culture de *E. coli*, et de la valider en simulation avant de l'implanter sur un vrai bioréacteur. Comme expliqué au chapitre 6, l'implantation initialement prévue sur un bioréacteur *E. coli* au laboratoire de biochimie de la Faculté Polytechnique de Mons n'a malheureusement pas pu être réalisée, suite à de nombreux problèmes au niveau de l'instrumentation du bioréacteur (en particulier le capteur d'acétate). Néanmoins, une prise en main de l'installation et la préparation de la pré-culture et des milieux de cultures ont été effectuées sur site. L'Annexe 4 présente le montage expérimental disponible au laboratoire de biochimie, utilisé pour réaliser plusieurs expériences en boucle ouverte, sans pouvoir finaliser l'implantation de la structure bouclée.

## 7.2 Apports scientifiques et originalité du travail

L'originalité du travail réside principalement dans l'application d'une des techniques de commande avancée la plus implantée dans un contexte industriel, en l'occurrence la commande prédictive (ici non-linéaire), au domaine des biotechnologies. Ce travail a donné lieu à plusieurs avancées scientifiques énumérées ci-dessous :

- Les modèles mathématiques obtenus à partir du schéma réactionnel restent complexes et difficiles à identifier. Ils font intervenir de nombreux paramètres. Sachant que l'identification paramétrique s'avère relativement complexe dans le cas des bioprocédés, simplifier le modèle (et donc restreindre le nombre de paramètres à identifier) est donc une démarche fondamentale, les erreurs de modélisation introduites devant être par la suite compensées par une loi de commande suffisamment robuste. En comparant les performances de la loi de commande prédictive non-linéaire à celles obtenues avec une loi de commande adaptative linéarisante et une loi de commande par modèle générique, il ressort que la loi de commande proposée a le grand avantage d'être applicable sur le modèle simplifié et de fournir de bonnes performances, meilleures que les deux autres commandes.
- Le principal apport au niveau de l'identification paramétrique repose sur la stratégie basée sur l'analyse de sensibilité. L'étude de sensibilité du modèle par rapport à ses paramètres permet de classer les paramètres selon leur influence pour chaque régime et/ou mode de fonctionnement. Seuls les paramètres les plus influents sont identifiés avec le scénario associé. Ceci permet d'éviter les problèmes d'identifiabilité et de mauvais conditionnement lors de l'identification paramétrique.
- Le profil optimal d'alimentation maximisant la croissance de la biomasse a pu être justifié de façon mathématique. En effet, le fonctionnement optimal en maintenant la concentration en acétate proche de zéro est un résultat connu dans la littérature dans la communauté biologiste. Cependant, à notre connaissance, la mise en évidence de l'optimalité de cette solution n'a pas été établie par le passé. La caractérisation analytique du profil optimal est également un résultat nouveau. D'autre part, la mise en évidence d'un point de fonctionnement optimal, caractérisé par un couple concentration en acétate/concentration en substrat est également novateur, et permet d'expliquer un certain nombre de travaux et résultats dans la littérature (non mis en évidence jusque là).
- La stratégie proposée pour l'application de la commande prédictive non-linéaire au bioprocédé *E. coli* fait apparaître plusieurs apports dans un contexte non-linéaire. Tout d'abord, la formulation du critère de la commande prédictive prend en compte deux écarts ; celui de la concentration en acétate par rapport à la concentration en acétate de référence et celui du débit d'alimentation (la commande) par rapport au profil de référence. De plus, pour éviter les problèmes liés à la résolution directe du problème d'optimisation sous contraintes de la commande prédictive non-linéaire, il a été proposé de transformer ce problème en un problème de programmation non-linéaire sans contraintes. La nouvelle formulation du problème d'optimisation est résolue à l'aide de l'approche séquentielle CVP, de sorte que la complexité numérique du calcul de la loi de commande s'en trouve considérablement réduite, ce qui facilitera une implantation de cette stratégie de commande sur un système réel.

- Le fait d'inclure la différence objet/modèle dans la structure de commande est un point déterminant dans l'objectif d'une implantation de la loi de commande développée sur un vrai bioréacteur. En effet, le modèle étant souvent « imparfait », surtout dans le cas des bioprocédés, il est indispensable de tenir compte de cette différence dans la structure de la loi de commande.

Enfin, il est à noter que la démarche proposée tout au long de cette thèse n'est pas spécifique au bioréacteur considéré et peut être appliquée à d'autres types de culture, quel que soit le mode de fonctionnement du bioréacteur, fed-batch, continu ou perfusé. Dans ce sens, la stratégie de commande développée dans ce mémoire a été appliquée avec succès à la commande d'un photo-bioréacteur de culture de microalgues *Porphyridium purpureum*, en mode continu [Becerra *et al.*, 2008]. Le but de la loi de commande dans ce cas est de maintenir la concentration en biomasse à une valeur donnée, en commandant le débit d'alimentation.

## 7.3 Perspectives

Ce travail de thèse, qui comme mentionné ci-dessus constitue une étude préliminaire à une implantation en milieu industriel, laisse apparaître de nombreuses perspectives à la fois niveau théorique et au niveau expérimental.

Au niveau théorique, trois points importants peuvent constituer une suite à ces travaux :

- La détermination du profil optimal maximisant la croissance de la biomasse. Une première piste serait de compléter l'étude déjà entreprise, en optimisant les instants de commutation entre les différents arcs optimaux. Dans ce cas, le scénario de succession des arcs peut ne pas être optimal, puisqu'il sera choisi a priori. Une deuxième piste consistera donc à déterminer le scénario optimal d'alimentation. Dans ce cas, des techniques d'optimisation sophistiquées doivent être utilisées pour résoudre le problème d'optimisation en considérant la structure « hybride » du modèle.
- La robustification de la loi de commande. la prise en compte de contraintes de robustesse dès la phase de synthèse de la loi de commande devra être approfondie. Une piste consiste à résoudre un problème d'optimisation de type min-max. Ceci permettra d'évaluer les performances en robustesse de la loi de commande prédictive non-linéaire de manière théorique.
- La synthèse d'observateur pour reconstituer l'état du système. Ce point peut s'avérer très important pour éviter les problèmes rencontrés au niveau de l'instrumentation du bioréacteur, en particulier en ce qui concerne la mesure de l'acétate. On pourrait ainsi envisager de ne mesurer que l'oxygène et le dioxyde de carbone dissous. S'affranchir de capteurs très onéreux est également très intéressant d'un point de vue économique.

Au niveau expérimental, l'implantation de la loi de commande proposée dans ce mémoire devra être réalisée tout d'abord sur un bioréacteur à échelle de laboratoire de culture de bactéries *E. coli*. Cette implantation se fera principalement en deux phases : Une phase préliminaire d'identification pour ajuster les valeurs des paramètres du modèle considéré, puis une phase d'implantation de la loi de commande et de validation.

La même démarche sera envisagée pour le cas de la commande de la culture de microalgues en mode continu, sur un photobioréacteur instrumenté du Laboratoire de Génie de Procédés et Matériaux de l'Ecole Centrale de Paris. Cette validation illustrera ainsi le caractère générique de la démarche mise en œuvre.



## 8. Références bibliographiques

**Andrews, J.** A mathematical model for the continuous culture of microorganism utilizing inhibitory substrate. *Biotechnology and Bioengineering*, 10, p. 707-723. 1968.

**Akesson, M., Hagander, P., Axelsson, P.** A probing feeding strategy for *Escherichia coli* cultures. *Biotechnology Techniques*. 13, 523-528. 1999.

**Alba, M.J.G., Calvo, G.** Characterization of bioreaction processes: aerobic *Escherichia coli* cultures. *Journal of Biotechnology*, 84, p. 107-118. 2000.

**Al Seyab, R.K., Cao, Y.** Nonlinear model predictive control for the ALSTOM gasifier. *Journal of Process control*, 16, p. 795-808. 2006.

**Balasubramhanya, L.S., Doyle III, F.J.** Nonlinear model-based control of a batch reactive distillation column. *Journal of Process Control*, 10, p. 209-218. 2000.

**Balsa-Canto, E., Banga, J.R., Alonso, A.A., Vassiliadis, V.S.** Dynamic optimization of chemical and biochemical processes using restricted second-order information. *Computers and Chemical Engineering*, 25, 4-6, 539-546. 2001.

**Banga, J.R., Alonso, A.A., Singh, R.P.** Stochastic dynamic optimization of batch and continuous bioprocesses. *Biotechnol. Prog.*, 13, 3, p. 326-335. 1997.

**Banga, J.R., Balsa-Canto, E., Moles, C.G., Alonso, A.A.** Dynamic optimization of bioprocesses: efficient and robust numerical strategies, *Journal of Biotechnology*, 4, p. 407-419. 2005.

**Bastin, G., Dochain, D.** On-line estimation and adaptive control of bioreactors. Elsevier Science Publishers. Amsterdam. 1990.

**Becerra-Celis, G., Hafidi, G., Tebbani, S., Dumur, D., Isambert, A.** Nonlinear predictive control for continuous microalgae cultivation process in a photobioreactor, 10<sup>th</sup> IEEE International Conference on Control, Automation, Robotics and Vision ICARCV 08, Hanoi, Vietnam, décembre 2008.

**Bequette, B.W.** Nonlinear control of chemical processes: A review, *Ind. Eng. Chem. Res*, 30, p. 1391-1413. 1991.

**Bernard, O.** Mass balance modelling of bioprocesses. A. Agrachev. (Ed). Int. center for theoretical physics, Trieste, 769-812. 2002.

**Biegler, L.T., Cervantes, A.M., Wachter, A.** Advances in simultaneous strategies for dynamic process optimization, *Chem. Eng. Science*, 47, p. 575-593. 2002.

**Bogaerts, Ph.** Contribution à la modélisation mathématique pour la simulation et l'observation d'états des bioprocédés. Belgique. 1999.

**Bonnans, J.F., Chartier, P., Zidani, H.** Discrete Approximations of the Hamilton-Jacobi Equation for an Optimal Control Problem of a Differential-Algebraic System, Rapport de recherche de l'INRIA. 2001.

**Bonvin, D., Srinivasan, B., Ruppen, D.** Dynamic optimization in the batch chemical industry. *Chemical Process Control*, 6. 2001.

- Boucher, P., Dumur, D.** La commande prédictive. Editions Technip, Paris. 1996.
- Bryson, A.E., Ho, Y.C.** Applied Optimal Control. Blaisdell Publishing Company. 1969.
- Bryson, A.E., Ho, Y.C.** Applied optimal control. John Wiley, New York. 1975.
- Camacho, E.F., Bordons, C.** Model Predictive Control in the process industry. Springer, London. 1998.
- Camacho, E.F., Bordons, C.** Model Predictive Control, Springer. 2003.
- Chen, L., Bastin, G.** Structural identifiability of the yield coefficients in bioprocesses when the reaction rates are unknown. *Mathematical Biosciences*, 132, p. 35-67. 1996.
- Chen, L.** Modelling, identifiability and control of complex biotechnological systems. Université Catholique de Louvain, Belgique, thèse de doctorat, 1992.
- Clarke, D.W., Mohtadi C., Tuffs, P.S.** Generalized Predictive Control : I. The Basic Algorithm. *Automatica*, 23 (2), 137. 1987.
- Clarke, D.W., Mohtadi C., Tuffs, P.S.** Generalized Predictive Control : II. Extensions and Interpretations. *Automatica*, 23 (2), 149. 1987.
- Cockshott, A.R., Bogle, I.D.L.** Modelling a recombinant *E. coli* fermentation producing bovin somatotropin. Modelling and control of Biotechnical Processes 1992 (2<sup>nd</sup> IFAC Symp. And 5<sup>th</sup> Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA). (Karim, M.N. and Stephanopoulos, G., Eds.). Pergamon, Oxford, p. 219-222. 1992.
- Company, O., Pierrot, F.** A new 3T-1R parallel robot. ICAR'99, Tokyo, Japan, October 25-27, p. 557-562. 1999.
- Costello, D.J., Lee, P.L., Greenfield, P.F.** Anaerobic digestion control by generic model control. *Bioprocess Engineering*, 4, p. 119-122. 1989.
- Cott, B.J., Durham, R.G., Lee, P.L. Sullivan, G.R.** Process model-based engineering. *Computers and chemical Engineering* 13, 9, p. 973-984. 1989.
- Cutler, C.R., Ramaker, B.L.** Dynamic matrix control- a computer control algorithm. AIChE Annual Meeting, Houston, Texas. 1979.
- Cutler, C.R., Ramaker, B.L.** Dynamic matrix control: a computer control algorithm. Joint Automatic Control Conference, San Francisco (CA). 1980.
- De Keyser, R.M.C.** A gentle introduction to model based predictive control. EC-PADI2 International Conference on Control Engineering and Signal Processing. Lima, Peru, Plenary Paper. 1998.
- De Mey, M., Lequeux, G.J., Beauprez, J.J., Maertens, Van Horen, E., Soetaert, W.K., Vanrolleghem, P.A., Vandamme, E.J.** Comparison of different strategies to reduce acetate formation in *Escherichia coli*. *Biotechnol. Prog.*, 23, 5, 1053-1063. 2007.
- Diaz, C., Dieu, P., Feuillerat, C., Lelong, P., Salomé, M.** Adaptive predictive control of dissolved oxygen concentration in a laboratory-scale bioreactor. *Journal of Biotechnology*, 43, 21-32. 1995.
- Diaz, C., Dieu, P., Feuillerat, C., Lelong, P., Salomé, M.** Simultaneous adaptive predictive control of the partial pressures of dissolved oxygen (pO<sub>2</sub>) and dissolved carbon dioxide (pCO<sub>2</sub>) in a laboratory-scale bioreactor. *Journal of Biotechnology*, 52, 135-150. 1996.

- Diehl, M., Gerhard, J., Marquardt, W., Mönnigmann, M.** Numerical solution approaches for robust nonlinear optimal control problems. *Computers and Chemical Engineering*, 32, 1279-1292. 2008.
- Diwekar, U.M.** *Batch Distillation: Simulation, Optimal Design and Control*. Taylor and Francis, Washington. 1995.
- Dochain, D., Vanrolleghem, P.** Identification de modèles de bioprocédés. In *Automatique des bioprocédés*, Hermes, chapitre 3. p. 53-82. 2001.
- Douglas, P.L., Fountain, P.S., Sullivan, G.R., Zhou, W.** Model-based control of a high-purity distillation column. *Canadian Journal of Chemical Engineering*, 72, 6, p. 1055-1065. 1994.
- Eitman, M.A., Altman, E.** Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol.* 24, 11, 530-533. 2006.
- Essen, H.V., Hijmeijer, H.** Nonlinear model predictive control of constrained mobile robots. In: *Proceedings of the European Control Conference*. Porto, Portugal. p. 1157-1162. 2001.
- Fikar, M., Latifi, M. A., Fournier, F., Creff, Y.** Control-vector parametrization versus iterative dynamic programming in dynamic optimization of a distillation column. *Comp. Chem. Eng.*, 22, p. 625-628. 1998.
- Fletcher, R.** *Practical methods of optimization* New York, Wiley. 1987.
- Fletcher, R.** *Practical methods of optimization*, Wiley-Interscience Publication, Chichester, ISBN 0471915475. 1991.
- Galvanauskas, V., Simutis, R., Volk, N., Lübbert, A.** Model based design of a biochemical cultivation process. *Bioprocess Engineering*, 18, p. 227-234. 1998.
- Garcia, C.E., Morari, M.** Internal Model Control. A unifying review and some new results. *I&EC Process Design and Development*, 21, 308-323. 1982
- Goh, C.J., Teo, K.L.** Control parameterization: A unified approach to optimal control problems with general constraints. *Automatica*, 24, p. 3-18. 1988.
- Hammersley, J.M., Handscomb, D.C.** *Les méthodes de Monte Carlo*. Dunod. 1967.
- Hicks, J., Mohan, A., Ray, W.H.** The optimal control of polymerization reactors. *Can. J. of Chem. Engng.*, p. 47-590. 1969.
- Huang, Y.C., Yeh, H.D.** The use of sensitivity analysis in on-line aquifer parameter estimation. *Journal of Hydrology*, 335, 406-418. 2007.
- Irving, E., Falinower, C.M., Fonte, C.** Adaptive Generalized Predictive Control with Multiple Reference Model, *Proceedings 2<sup>nd</sup> ACASP/86*, Lund, Juin, 1986.
- Jaspan, R.K., Coull, J.** Trajectory optimization techniques in chemical reaction engineering. II. Comparison of the methods. *AIChE J.*, 18(4), p. 867-869. 1972.
- Jenzsch, M., Simutis, R., Luebbert, A.** Generic model control of the specific growth rate in recombinant *Escherichia coli* cultivations. *Journal of Biotechnology* 122, p. 483-493. 2006.
- Kawathekar, R., Riggs, J.B.** Nonlinear model predictive control of a reactive distillation column. *Control Engineering Practice* 15, p. 231-239. 2007.

**Keerthi, S.S., Gilbert, E.G.** Optimal, infinite horizon feedback laws for general class of constrained discrete time systems. Stability and Moving horizon approximations. Journal of optimization theory and applications, 57, 265-293. 1988.

**Kleman, G.L., Strohl, W.R.** Acetate Metabolism by *Escherichia coli* in High Cell Density Fermentation. Appl. Environ. Microbiol. 57, p. 3952-3958. 1994.

**Ko, Y.F., Bentley, W.E, Weigand, W.A.** An integrated metabolic modelling approach to describe the energy efficiency of *Escherichia coli* fermentations under oxygen-limited conditions: Cellular energetics, carbon flux, and acetate production. Biotechnology and Bioengineering. 42, p. 843-853. 1993.

**Ko, Y.F., Bentley, W.E, Weigand, W.A.** A metabolic model of cellular energetics and carbon flux during aerobic *Escherichia coli* fermentation. Biotechnology and Bioengineering. 43, p.847-855. 1994.

**Ko, C.L., Wang, F.S.** Run-to-run fed-batch optimization for protein production using recombinant *Escherichia coli*, Biochemical Engineering Journal 30, 3, 279-285, 25 June 2006.

**Korz, D.J., Rinas, U., Hellmuth, K., Sanders, E.A., Deckwer, W.D.** Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. Journal of Biotechnology, 39, 59-65. 1995.

**Lee, E.B., Markus, L.** Foundations of optimal control theory, John Wiley and Sons, New York. 1967.

**Lee, E.S.** Quasi linearization and invariant imbedding with applications to chemical engineering and adaptive control, Academic Press, New York. 1968.

**Lee, P.L., Sullivan, G.R.** Generic model control (GMC). Computers and Chemical Engineering 12, 6, p. 573-580. 1988.

**Ljung, L.** System Identification-theory for the user Prentice Hall. Englewood Cliffs, New Jersey. 1987.

**Luenberger, D.G.** Optimization by vector space methods, John Wiley and sons, New York. 1969.

**Luli, G.W., Strohl, W.R.** Comparison of growth, acetate production and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. Appl. Environ. Microbiol. 56, 1004-1011. 1990.

**Maciejowski, J.** Predictive Control with Constraints, Prentice Hall. 2002.

**Mayne, D.Q., Michalska, H.** Receding-horizon control of nonlinear systems», IEEE transactions on automatic control, 35, 5, 814-824. 1990.

**Mayne D.Q., Rawlings, J.B., Rao, C.V., Sokaert, P.O.** Constrained model predictive control : stability and optimality. Automatica, 36, 789-814. 2000.

**Modak, J., Lim, H., Tayeb, Y.** General characteristics of optimal feed rate profiles for various fed-batch fermentation processes. Biotechnology and Bioengineering 28, p. 1396-1407. 1986.

**Monod, J.** Recherches sur la croissance des cultures bactériennes. Paris, France, Hermes. 1942.

- Morari, M., Lee, J.H.** Model predictive control: past, present and future. *Computers and Chemical Engineering*, 23, p. 667-682. 1999.
- M'Saad, M.** Sur l'applicabilité de la commande adaptative. Thèse d'état de l'Institut Polytechnique de Grenoble, Mars 1987.
- Munack, A.** Optimization and sampling. Measurements in Bioreactor Systems. 2<sup>nd</sup> Ed. (Rehm, H. J., Reed, G., Pühler, A., and Stadler, P., Eds.). In series: Biotechnology (vol. 2). VCH, Weinheim, p. 252-264. 1991.
- Nagy, Z.K.** Model based control of a yeast fermentation bioreactor using optimally designed artificial neural networks. *Chemical Engineering Journal*, 127, p. 95-109. 2007.
- Nakano, K., Rischke, M., Sato, S., Maerkl, H.** Influence of acetic acid on the growth of *Escherichia coli* K12 during high-cell-density cultivation in a dialysis reactor. *Appl. Microbiol. Biotechnol.* 48, 5, 597-601. 1997.
- Pertev, C., Türker, M., Berber, R.** Dynamic modelling, sensitivity analysis and parameter estimation of industrial yeast fermenters. *Computers and Chemical Engineering*. 21, p. 739-744. 1997.
- Queinnec, I., Dahhou, B., M'Saad, M.** On adaptive control of fed-batch fermentation processes, *International Journal of Adaptive Control and Signal Processing*, 6, 6, p. 521-536. 1992.
- Ramaswamy, S., Cutright, T.J., Qammar, H.K.** Control of a continuous bioreactor using model predictive control. *Process Biochemistry*, 40, p. 2763-2770. 2005.
- Ray, W.H.** *Advanced Process Control*. McGraw-Hill, New York. 1981.
- Renard, F.** Commande robuste de bioprocédés opérés en mode fed-batch – Application industrielle à des cultures de *Saccharomyces Cerevisiae*. Thèse de doctorat. Mons, Belgique. 2005.
- Renfro, J.G., Morshedi, A.M., Asbjornsen.** Simultaneous optimization and solution of systems described by differential/algebraic equations. *Comp. Chem. Eng.*, 11, 5, p. 503-517. 1987.
- Richalet, J., Abu El Ata, S., Arber, C., Kuntze, M.B., Jacubasch, A., Schill, W.** Predictive Functional Control. Application to Fast and Accurate Robots. Proceedings 10<sup>th</sup> IFAC World Congress, Munich. July. 1987.
- Richalet, J., Rault, A., Testud, J.L., Papon, J.** Algorithmic control of industrial processes. In Proceedings of the 4th IFAC Symposium on identification and system parameter estimation, p. 1119-1167. 1976.
- Richalet, J., Rault, A., Testud, J.L., Papon, J.** Model predictive heuristic control : Applications to industrial processes. *Automatica*, 14, p. 413-428. 1978.
- Robinson, J.A.** Determining microbial parameters using nonlinear regression analysis: Advantages and limitations in microbial ecology. *Adv. Microb. Ecol.*, 8, p. 61-114. 1985.
- Rocha, I.** Model-based strategies for computer-aided operation of a recombinant *E. coli* fermentation, thèse de doctorat, Univesidade do Minho, Portugal. 2003.
- Rocha, M., Neves, J., Rocha, I., Ferreira, E.C.** Evolutionary algorithms for optimal control in fed-batch fermentation processes. *EvoWorkshop 2004, LNCS 3005*, 84-93. 2004.

- Rodrigues, J.A.D, Filho, R.M.** Production optimization with operating constraints for a fed-batch reactor with DMC predictive control. *Chemical Engineering Science*, 54, p. 2745-2751. 1999.
- Rossiter, J. A.** Model – based predictive control. A practical approach. CRC Press, ISBN 0-8493-1291-4. 2003.
- Sacks, M.E., Lee, S., Biesenberger, J.A.** Optimum policies for batch, chain addition polymerizations. *Chemical. Engineering. Science.*, p. 27-2281. 1972.
- Schlegel, M., Marquardt, W.** Detection and exploitation of the control switching structure in the solution of dynamic optimization problems. *Journal of Process Control*, 16, 3, 275-290. 2006.
- Schwarz, H.R.** Numerical Analysis- A Comprehensive Introduction. John Wiley, New York. 1989.
- Serban, R., Hindmarsh, A.C.** CVodes, the sensitivity-enabled ode solver in SUNDIALS, In Proceedings of the 5th International Conference on Multibody Systems, Nonlinear Dynamics and Control, Long Beach, CA. 2005. ASME.
- Smets, I., Bernaerts, K., Sun, J., Marchal, K., Vanderleyden, J., Van Impe, J.F.** Sensitivity function-based model reduction. A bacterial gene expression case study. *Biotechnology and Bioengineering*. 80, p. 195-200. 2002.
- Smets, I., Claes, J., November, E., Bastin, G., Van Impe, J.** Optimal adaptive control of (bio)chemical reactors: past, present and future. *Journal of process control*, 14, p. 795-805, 2004.
- Smets, I., Perrier, M., Bastin, G., Van Impe, J.** Commande optimale des bioréacteurs. Dans *Automatique des bioprocédés*, Hermes, chapitre 7, p. 181-212. 2001.
- Sorensen, E., Macchietto, S., Stuart, G., Skogestad, S.** Optimal control and on-line operation of reactive batch distillation. *Comp. Chem. Eng.*, 20, 12, p. 1491-1498. 1996.
- Srinivasan, B., Palanki, S., Bonvin, D.** Dynamic optimization of batch processes: I. Characterization of the nominal solution. *Computers and Chemical Engineering*, 27, Issue 1, 15, p. 1-26. 2003. Best Paper of the Year 2003 Award.
- Sussmann, H.J., Jurdjevic, V.** Controllability of nonlinear systems. *Journal of Differential Equations*, 12, p. 95-16. 1972.
- Tebbani, S., Dumur, D., Hafidi, G.** Optimisation du fonctionnement d'un bioréacteur en mode fed-batch. *Optimisation et commande en génie des procédés SIMO 06*, Toulouse, octobre 2006.
- Vassiliadis, V.S.** Computational solution of dynamic optimization problems with general differential-algebraic constraints. PhD thesis, Imperial College, University of London, London, UK. 1993.
- Xaumier, F., Le Lann, M.V. Cabassud, M., Casamatta, G.** Experimental application of nonlinear model predictive control : temperature control of an industrial semi-batch pilot-plant reactor. *Journal of Process Control*, 12, p. 687-693. 2002.
- Xu, B., Jahic, M., Enfors, S.O.** Modeling of overflow metabolism in batch and fed batch cultures of *Escherichia coli*. *Biotechnology Progress*, 15, p. 81-90. 1999.

**Zhu, G.Y., Zamamiri, A., Henson, M.A., Hjortso, M.A.** Model predictive control of continuous yeast bioreactors using cell population balance models. *Chemical Engineering Science*, 55, p. 6155-6167. 2000.



## 9. Annexes

### 9.1 Annexe 1 : Analyse de sensibilité en régime oxydatif

Cette annexe présente l'analyse de sensibilité du modèle de bioréacteur *E. coli* en régime oxydatif (la même étude est présentée au chapitre 3 dans le cas du régime oxydo-fermentatif). Rappelons que, lors du fonctionnement en régime oxydatif, il est nécessaire de distinguer deux cas, comme suit:

- le « sous-régime 1 » qui est activé lorsque la somme des taux de croissance spécifiques utilisés pour l'oxydation de glucose et de l'acétate est inférieure à la capacité oxydative, c'est-à-dire  $(q_{OS} + q_{OA}) \leq q_{Scrit} k_{os}$ ,
- le « sous-régime 2 » qui est activé lorsque la somme des taux de croissance spécifiques utilisés pour l'oxydation de glucose et de l'acétate est supérieure à la capacité oxydative, c'est-à-dire  $(q_{OS} + q_{OA}) \geq q_{Scrit} k_{os}$ . Ce sous-régime assure la continuité et la transition entre le régime oxydo-fermentatif et le régime oxydatif.

Dans ce qui suit, l'analyse de sensibilité est donc présentée pour chaque sous-régime en tenant compte des deux modes d'alimentation (batch et fed-batch).

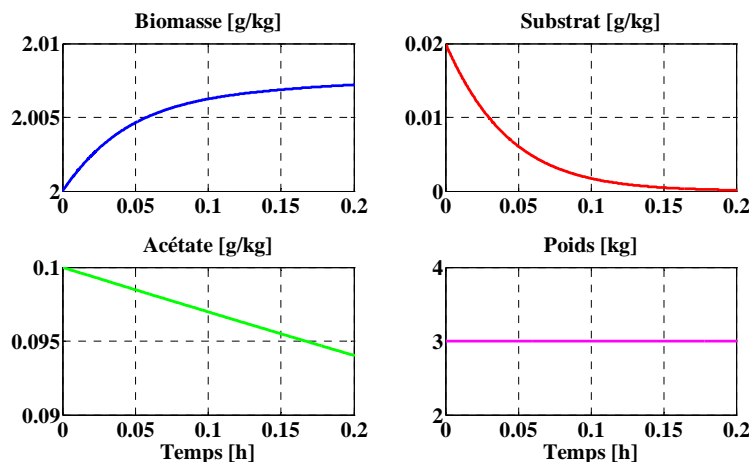
#### 9.1.1 Fonctionnement en sous-régime 1

Cette partie présente l'analyse de sensibilité du modèle aux paramètres pour un fonctionnement en régime oxydatif et en sous-régime 1 et pour les deux modes d'alimentation (batch et fed-batch).

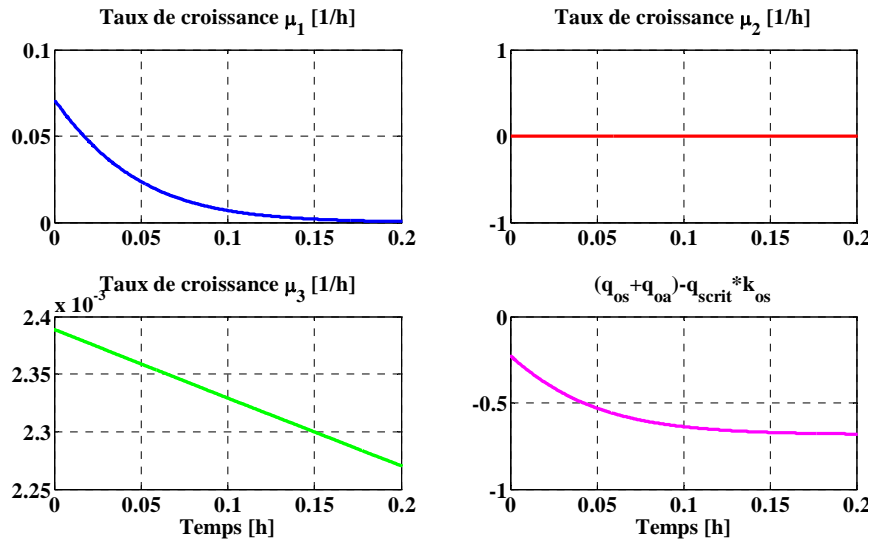
##### 9.1.1.1 Bioréacteur en mode batch

Pour un fonctionnement en régime oxydatif, en sous-régime 1 et en mode batch, les conditions choisies sont :

- le vecteur initial  $[X_0 \ S_0 \ A_0 \ W_0] = [2 \text{ g/kg} \ 0,02 \text{ g/kg} \ 0,1 \text{ g/kg} \ 3 \text{ kg}]$ ,
- l'intervalle d'analyse est restreint à  $[0 \ 0,2 \text{ h}]$ .



**Figure 9.1.** Evolutions des variables d'état en mode batch, en régime oxydatif et en sous-régime 1 sur l'intervalle  $[0 \ 0,2 \text{ h}]$ .

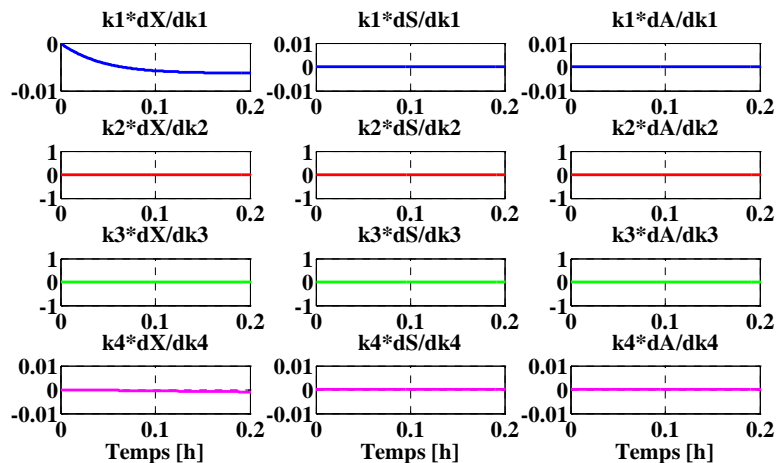


**Figure 9.2.** Evolutions des taux de croissance en mode batch, en régime oxydatif et en sous-régime 1 sur l'intervalle  $[0, 0,2 h]$ .

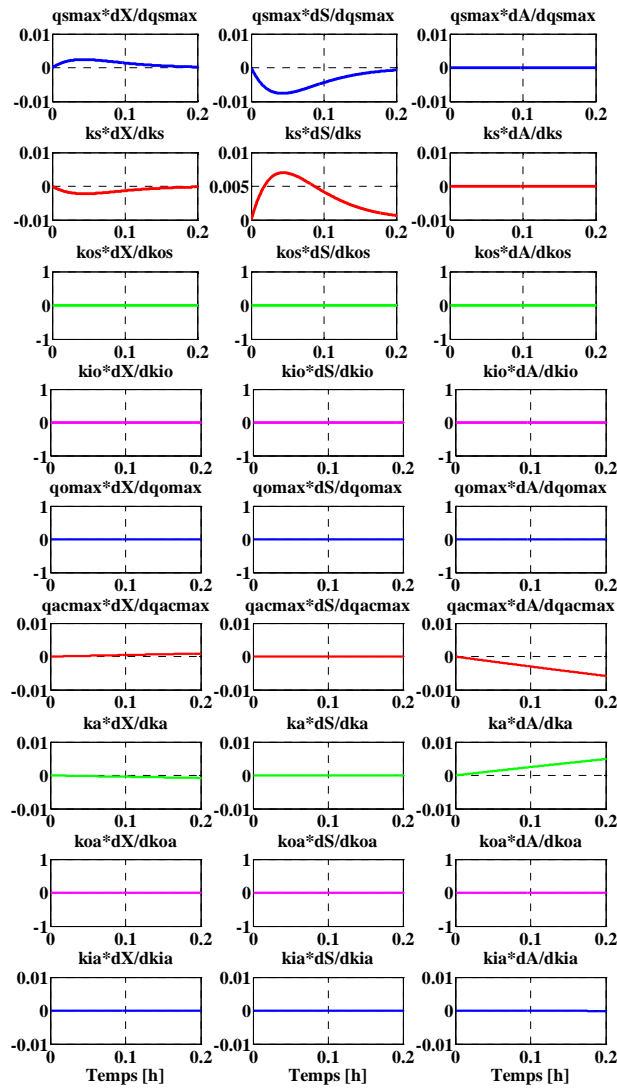
Les Figures 9.1 et 9.2 représentent les évolutions des variables d'état et des taux de croissance. Ces figures montrent que le cas étudié est un fonctionnement en mode batch (poids constant), en régime oxydatif ( $\mu_2 = 0$ ) et en sous-régime 1 ( $(q_{OS} + q_{OA}) - q_{Scrit} k_{os} \leq 0$ ).

Les Figures 9.3 et 9.4 représentent les fonctions de sensibilité normalisées des coefficients stœchiométriques et des coefficients cinétiques. Pour le régime considéré, les paramètres  $k_2$ ,  $k_3$ ,  $k_{oa}$ ,  $q_{Omax}$  et  $k_{os}$  n'interviennent pas.

Le Tableau 9.1 présente le classement des huit autres paramètres selon leur influence pour chaque variable d'état.



**Figure 9.3.** Fonctions de sensibilité normalisées des coefficients stœchiométriques pour chaque variable d'état en mode batch, en régime oxydatif et en sous-régime 1.



**Figure 9.4.** Fonctions de sensibilité normalisées des coefficients cinétiques pour chaque variable d'état en mode batch, en régime oxydatif et en sous-régime 1.

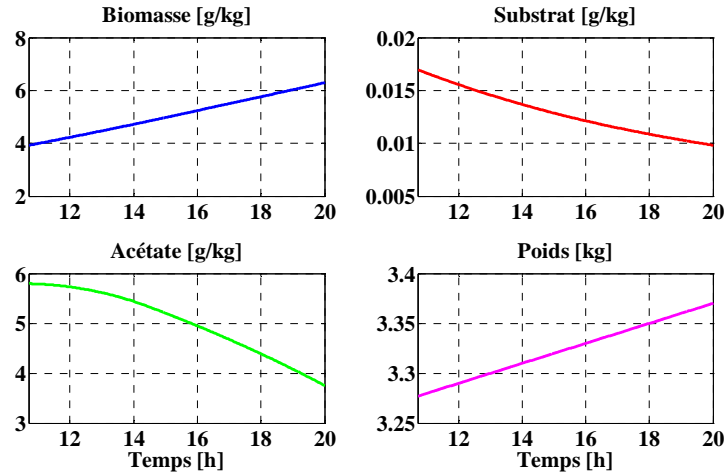
Influence sur la biomasse $X$	Influence sur le substrat $S$	Influence sur l'acétate $A$
$k_1$		
$k_4$		
$q_{ac\ max}$		$q_{ac\ max}$
$k_a$	$q_{s\ max}$	$k_a$
$q_{s\ max}$	$k_s$	$k_{ia}$
$k_s$		
$k_{ia}$		

**Tableau 9. 1.** Classement des paramètres selon leur influence pour chaque variable d'état (du plus influent au moins influent) en mode batch, en régime oxydatif et en sous-régime 1.

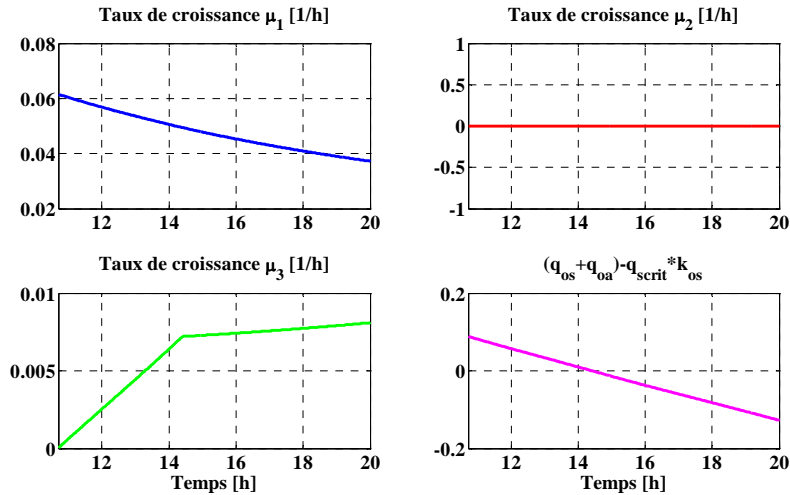
### 9.1.1.2 Bioréacteur en mode fed-batch

Le scénario choisi pour le fonctionnement en mode fed-batch et en régime oxydatif est tel que :

- le vecteur initial  $[X_0 \ S_0 \ A_0 \ W_0] = [2 \text{ g/kg} \ 0,1 \text{ g/kg} \ 5 \text{ g/kg} \ 3,17 \text{ kg}]$ ,
- l'intervalle d'analyse est restreint à  $[10,7 \text{ h} \ 20 \text{ h}]$ ,
- le débit d'alimentation en glucose est constant égal à  $0,01 \text{ kg/h}$ .



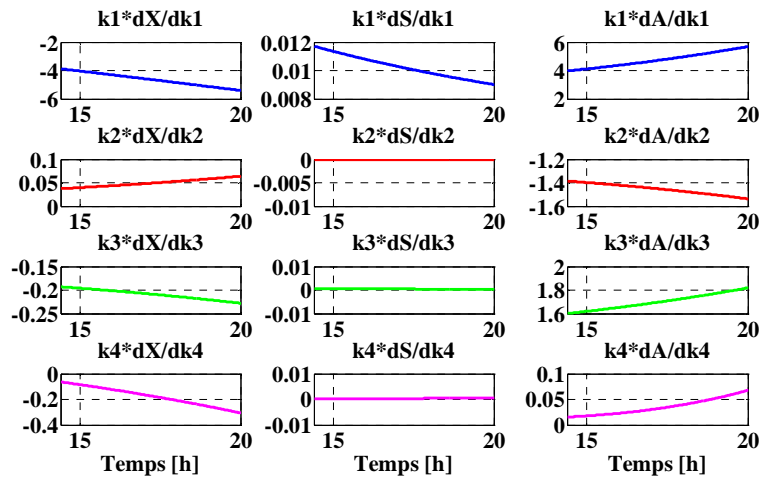
**Figure 9.5.** Evolutions des variables d'état du modèle en mode fed-batch et en régime oxydatif sur l'intervalle  $[10,7 \text{ h} \ 20 \text{ h}]$ .



**Figure 9.6.** Evolutions des taux de croissance en mode fed-batch et en régime oxydatif sur l'intervalle  $[10,7 \text{ h} \ 20 \text{ h}]$ .

Les Figures 9.5 et 9.6 représentent les évolutions des variables d'état et des taux de croissance. La Figure 9.6 montre que sur l'intervalle  $[10,7 \text{ h} \ 20 \text{ h}]$ , le bioréacteur fonctionne en régime oxydatif ( $\mu_2 = 0$ ) et que la différence  $(q_{OS} + q_{OA}) - q_{Scrit}^* k_{os}$  change de signe sur l'intervalle  $[10,7 \text{ h} \ 20 \text{ h}]$ . En effet les deux sous-régimes apparaissent dans ce cas, le premier sous-régime sur l'intervalle  $[14,4 \text{ h} \ 20 \text{ h}]$  et le second sur l'intervalle  $[10,7 \text{ h} \ 14,4 \text{ h}]$ . On se limite donc dans cette partie à l'étude de la sensibilité du modèle aux paramètres sur l'intervalle  $[14,4 \text{ h} \ 20 \text{ h}]$  (sous-régime 1).

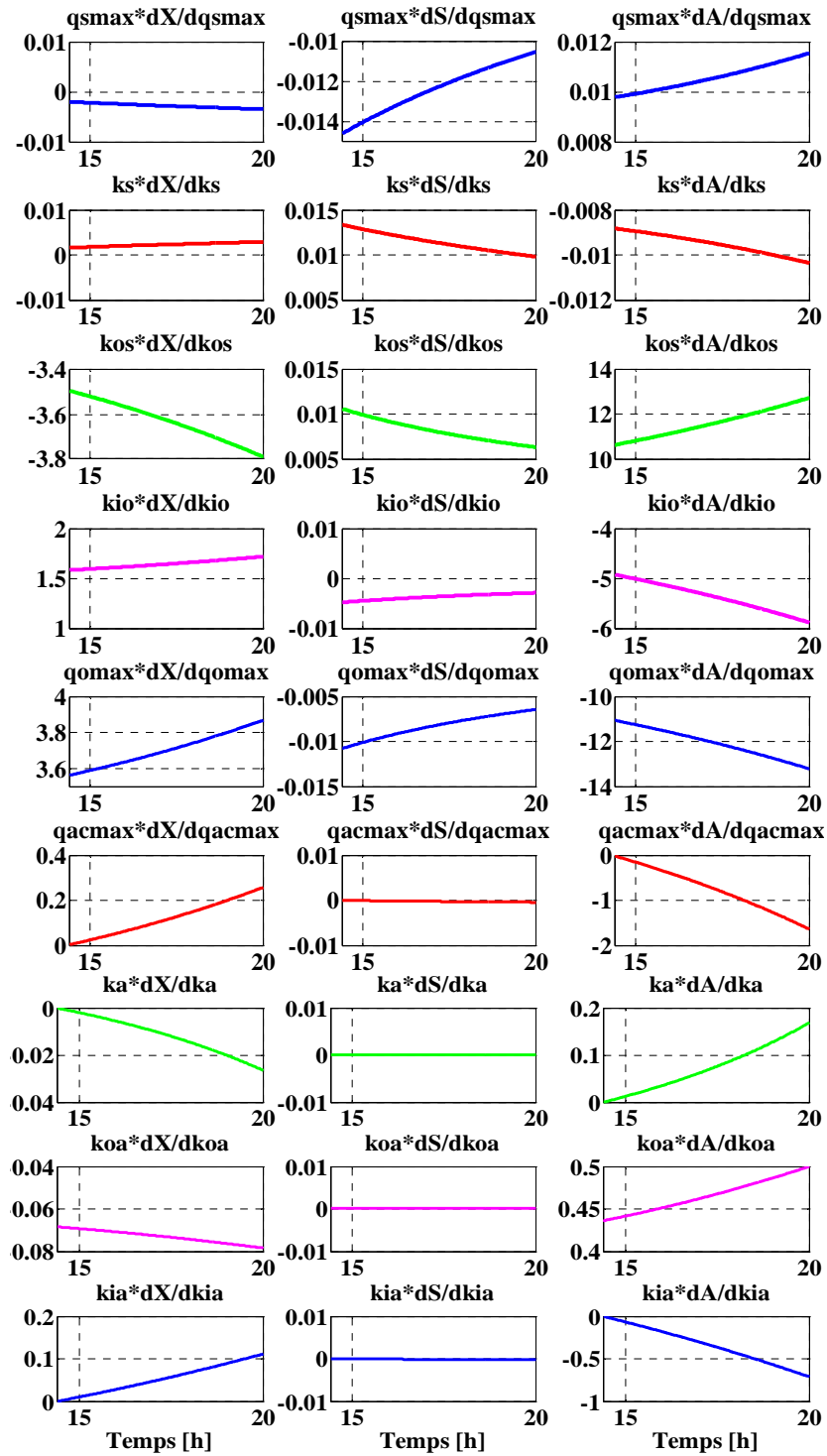
Les Figures 9.7 et 9.8 représentent les fonctions de sensibilité normalisées des coefficients stœchiométriques et des coefficients cinétiques en mode fed-batch, en régime oxydatif et en sous-régime 1. Ces figures montrent que tous les paramètres ont une influence sur les états, à différents degrés. Ces paramètres sont classés dans le Tableau 9.2 selon leur influence.



**Figure 9.7.** Fonctions de sensibilité normalisées des coefficients stœchiométriques pour chaque variable d'état en mode fed-batch, en régime oxydatif et en sous-régime 1.

Influence sur la biomasse $X$	Influence sur le substrat $S$	Influence sur l'acétate $A$
$k_1$	$k_s$	$k_{os}$
$k_{os}$	$k_1$	$q_{o \max}$
$q_{o \max}$	$q_{s \max}$	$k_{io}$
$k_{io}$	$k_{os}$	$k_1$
$k_4$	$q_{o \max}$	$k_3$
$k_3$	$k_{io}$	$q_{ac \max}$
$q_{ac \max}$	$k_4$	$k_2$
$k_{ia}$	$k_3$	$k_{ia}$
$k_2$	$q_{ac \max}$	$k_{oa}$
$k_{oa}$	$k_{oa}$	$k_a$
$k_a$	$k_{ia}$	$k_4$
$k_s$	$k_2$	$q_{s \max}$
$q_{s \max}$	$k_a$	$k_s$

**Tableau 9.2.** Classement des paramètres selon leur influence pour chaque variable d'état (du plus influent au moins influent) en mode fed-batch, en régime oxydatif et en sous-régime 1.



**Figure 9.8.** Fonctions de sensibilité normalisées des coefficients cinétiques pour chaque variable d'état en mode fed-batch, en régime oxydatif et en sous-régime 1.

### 9.1.2 Fonctionnement en sous-régime 2

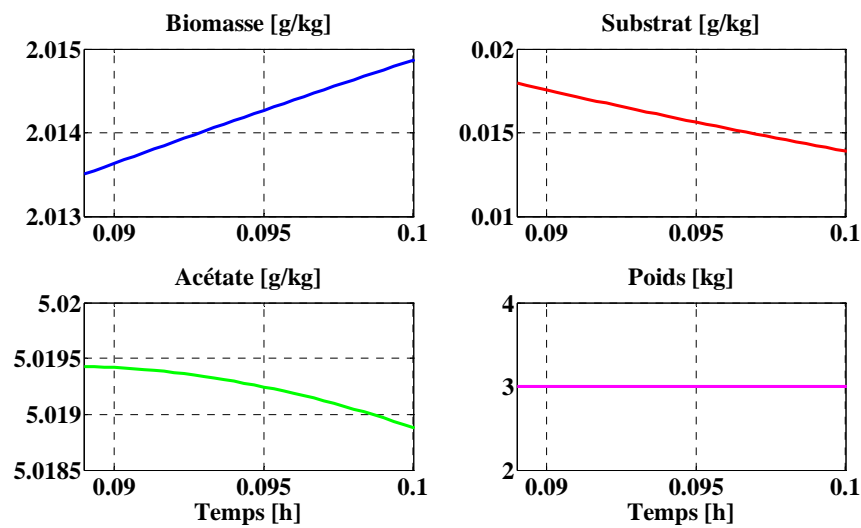
De la même manière que pour le fonctionnement en sous-régime 1, cette partie présente l'étude de sensibilité du modèle aux paramètres pour un fonctionnement en régime oxydatif et en sous-régime 2 et pour les deux modes d'alimentation (batch et fed-batch).

### 9.1.2.1 Bioréacteur en mode batch

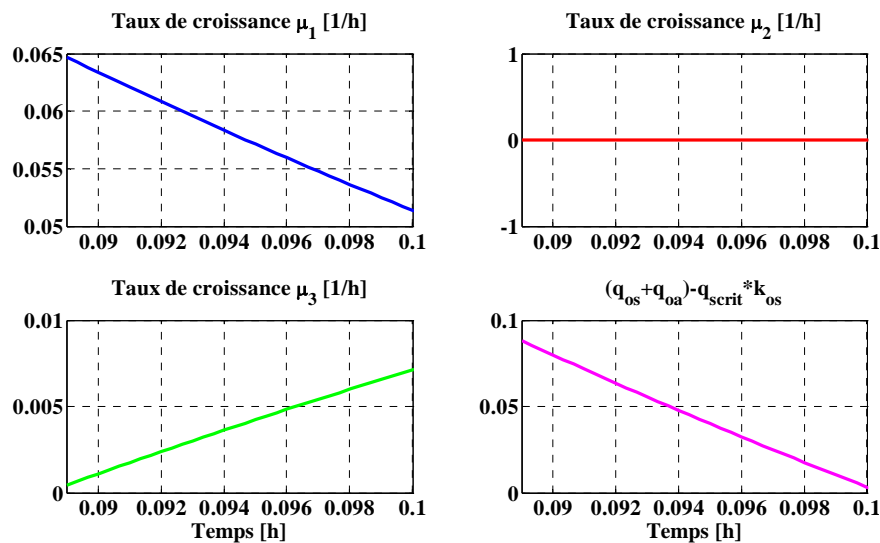
Pour se placer dans ce sous régime, le scénario virtuel choisi est le suivant :

- le vecteur initial  $[X_0 \ S_0 \ A_0 \ W_0] = [2 \text{ g/kg} \ 0,1 \text{ g/kg} \ 5 \text{ g/kg} \ 3 \text{ kg}]$ ,
- l'intervalle d'analyse est restreint à  $[0,089 \text{ h} \ 0,1 \text{ h}]$ .

Les Figures 9.9 et 9.10 représentent les évolutions des concentrations en biomasse, en substrat et en acétate et du poids, et les évolutions des taux de croissance. Le poids étant constant confirme le mode de fonctionnement du bioréacteur (Figure 9.9). Le taux de croissance  $\mu_2$  est nul, ce qui correspond au régime oxydatif (chapitre 2) et la différence  $(q_{OS} + q_{OA}) - q_{Scrit} k_{os}$  est positive sur l'intervalle considéré, donc le sous-régime 2 est actif.

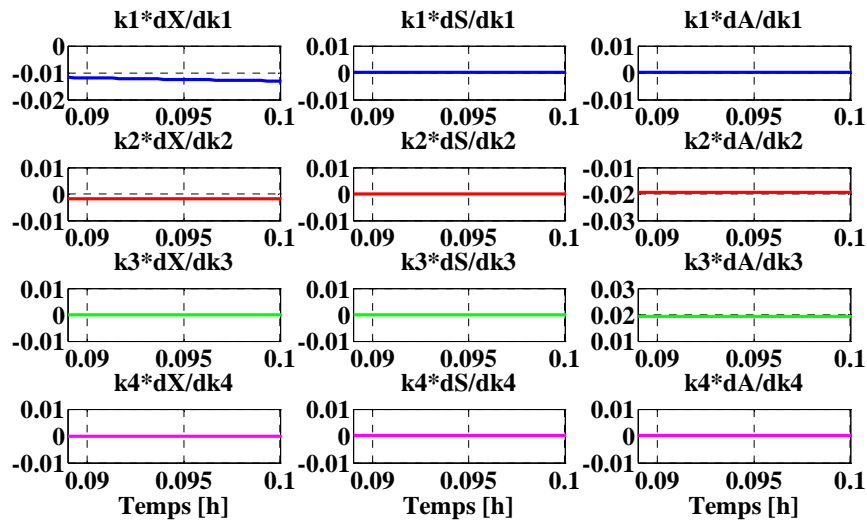


**Figure 9.9.** Evolutions des variables d'état du modèle en mode batch, en régime oxydatif et en sous-régime 2 et sur l'intervalle  $[0,089 \text{ h} \ 0,1 \text{ h}]$ .



**Figure 9.10.** Evolutions des taux de croissance en mode batch, en régime oxydatif et en sous-régime 2 et sur l'intervalle  $[0,089 \text{ h} \ 0,1 \text{ h}]$ .

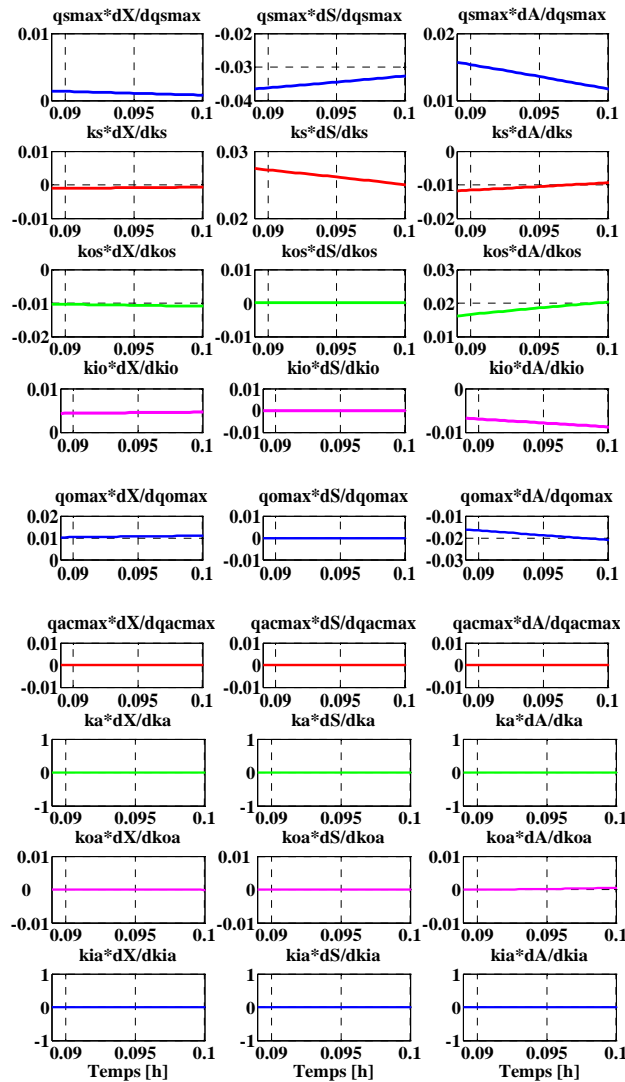
Les Figures 9.11 et 9.12 représentent les fonctions de sensibilité normalisées des coefficients stœchiométriques et des coefficients cinétiques. Les paramètres  $q_{ac\max}$ ,  $k_a$  et  $k_{ia}$  n'interviennent pas dans le cas de ce sous-régime. L'influence des dix paramètres restant sur le modèle pour chaque variable d'état est présentée dans le Tableau 9.3.



**Figure 9.11.** Fonctions de sensibilité normalisées des coefficients stœchiométriques pour chaque variable d'état en mode batch, en régime oxydatif et en sous-régime 2.

Influence sur la biomasse $X$	Influence sur le substrat $S$	Influence sur l'acétate $A$
$k_1$		$q_o \max$
$q_o \max$	$q_s \max$	$k_{os}$
$k_{os}$	$k_s$	$k_2$
$k_{io}$	$k_1$	$k_3$
$k_2$	$q_o \max$	$q_s \max$
$q_s \max$	$k_{os}$	$k_s$
$k_s$	$k_{io}$	$k_{io}$
$k_4$	$k_2$	$k_{oa}$
$k_{oa}$		

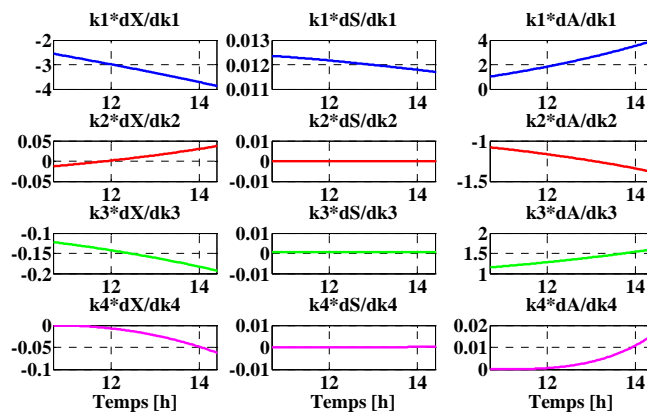
**Tableau 9.3.** Classement des paramètres selon leur influence pour chaque variable d'état (du plus influent au moins influent) en mode batch, en régime oxydatif et en sous-régime 2.



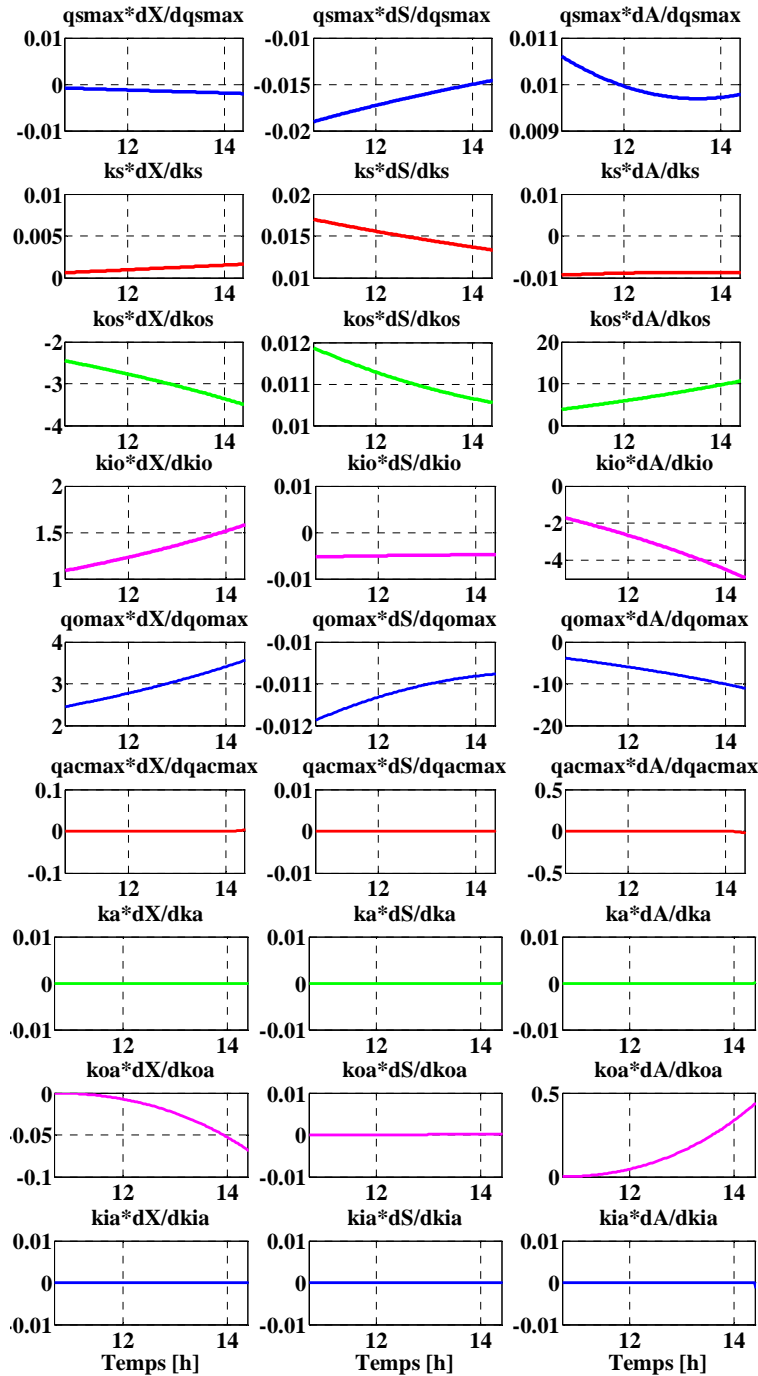
**Figure 9.12.** Fonctions de sensibilité normalisées des coefficients cinétiques pour chaque variable d'état en mode batch, en régime oxydatif et en sous-régime 2.

### 9.1.2.2 Bioréacteur en mode fed-batch

Le scénario choisi ici est le même que dans le cas du sous-régime 1 pour le mode d'alimentation fed-batch, sauf que l'on s'intéresse à l'intervalle [10,7h - 14,4h].



**Figure 9.13.** Fonctions de sensibilité normalisées des coefficients stœchiométriques pour chaque variable d'état en mode fed-batch, en régime oxydatif et en sous-régime 2.



**Figure 9.14.** Fonctions de sensibilité normalisées des coefficients cinétiques pour chaque variable d'état en mode fed-batch, en régime oxydatif et en sous-régime 2.

Les Figures 9.13 et 9.14 représentent les fonctions de sensibilité des coefficients stœchiométriques et des coefficients cinétiques pour chaque variable d'état. Comme pour le fonctionnement en mode batch, les paramètres  $q_{acmax}$ ,  $k_a$  et  $k_{ia}$  n'interviennent pas dans ce cas, donc n'ont aucune influence sur le modèle comme le montre la Figure 9.14.

Le Tableau 9.4 présente le classement des autres paramètres selon leur influence pour chaque variable d'état.

Influence sur la biomasse $X$	Influence sur le substrat $S$	Influence sur l'acétate $A$
$k_1$	$q_{s \max}$	$k_{os}$
$q_{o \max}$	$k_s$	$q_{o \max}$
$k_{os}$	$k_1$	$k_{io}$
$k_{io}$	$k_{os}$	$k_1$
$k_3$	$q_{o \max}$	$k_3$
$k_4$	$k_{io}$	$k_2$
$k_{oa}$	$k_3$	$k_{oa}$
$k_2$	$k_4$	$q_{ac \max}$
$q_{ac \max}$	$k_{oa}$	$k_{ia}$
$k_{ia}$	$k_2$	$k_a$
$k_a$	$q_{ac \max}$	$k_4$
$q_{s \max}$	$k_{ia}$	$q_{s \max}$
$k_s$	$k_a$	$k_s$

**Tableau 9.4.** Classement des paramètres selon leur influence pour chaque variable d'état (du plus influent au moins influent) en mode fed-batch, en régime oxydatif et en sous-régime 2.

### 9.1.3 Conclusions

On constate après l'étude de tous les régimes que l'influence des paramètres dépend des régimes, sous-régimes et modes d'alimentation. En effet, les Tableaux (9.1, 9.2, 9.3, 9.4) montrent que le classement des paramètres selon leur influence dépend des conditions opératoires.

La deuxième étape consiste comme pour l'autre régime à déterminer les dépendances linéaires entre les paramètres, en étudiant les matrices de Gram.

## 9.2 Annexe 2 : Expression détaillée de la matrice $M$

Le but de cette annexe est de préciser de façon plus détaillée l'expression de la matrice  $M$  présentée au chapitre 4, lors de la caractérisation de la solution optimale du problème de maximisation de la croissance de la biomasse dans le bioréacteur en fed-batch.

Rappelons que cette matrice s'écrit sous la forme :

$$M = [g \quad q \quad G] \quad (\text{A2.1})$$

$$\text{où: } g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad q = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g \quad \text{et} \quad G = \left( \frac{\partial q}{\partial x} f - \frac{\partial f}{\partial x} q \right) + \left( \frac{\partial q}{\partial x} g - \frac{\partial g}{\partial x} q \right).$$

La fonction  $f$  s'écrit sous la forme :

$$f = \begin{bmatrix} f_1 \\ f_2 \\ 0 \end{bmatrix} = \begin{bmatrix} r_X XW \\ r_A XW \\ 0 \end{bmatrix} \quad (\text{A2.2})$$

On pose dans ce qui suit :  $x_1 = XW$ ,  $x_2 = AW$ ,  $x_3 = W$ . Comme nous l'avons montré, les expressions des taux  $r_X$  et  $r_A$  dépendent du régime et/ ou sous-régime de fonctionnement. L'expression analytique de la matrice  $M$  obtenue grâce à la *Symbolic Matlab* est très longue, on donne ici uniquement les expressions des dérivées partielles de  $f$  selon le cas de fonctionnement.

La matrice des dérivées partielles de  $f$  s'écrit sous la forme :

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{A2.3})$$

- Régime oxydatif :

Pour ce régime de fonctionnement, les expressions des taux  $r_X$  et  $r_A$  sont différentes selon le sous-régime de fonctionnement, les expressions des éléments de la matrice  $\partial f / \partial x$  sont alors données pour chaque cas.

- Sous-régime 1 :

Rappelons que les expressions des taux  $r_X$  et  $r_A$  sont :

$$r_X = \frac{q_S}{k_1} + \frac{q_{AC}}{k_4} \quad \text{et} \quad r_A = -q_{AC} \quad (\text{A2.4})$$

Les expressions des éléments de la matrice  $\partial f / \partial x$  sont :

$$\frac{\partial f_1}{\partial x_1} = q_{s \max} k_1 / k_s + q_{ac \max} x_2 (k_a + x_2 / x_3) (k_{ia} + x_2 / x_3) / x_3 k_{ia} k_4$$

$$\frac{\partial f_2}{\partial x_1} = -q_{ac \max} x_2 (k_a + x_2 / x_3) (k_{ia} + x_2 / x_3) / x_3 k_{ia}$$

$$\frac{\partial f_1}{\partial x_2} = (q_{ac \max} k_{ia} k_4 (k_a + x_2 / x_3) / x_3 (k_{ia} + x_2 / x_3) - q_{ac \max} k_{ia} k_4 x_2 (k_a + x_2 / x_3)^2 / x_3^2 (k_{ia} + x_2 / x_3)^2) x_1$$

$$\frac{\partial f_2}{\partial x_2} = -q_{ac \max} k_{ia} (k_a + x_2 / x_3) / x_3 (k_{ia} + x_2 / x_3) x_1 + q_{ac \max} k_{ia} x_2 (k_a + x_2 / x_3)^2 / (k_{ia} + x_2 / x_3) x_1 x_3^2$$

$$+ q_{ac \max} k_{ia} (k_a + x_2 / x_3) x_2 / (k_{ia} + x_2 / x_3)^2 x_3^2 x_1$$

$$\frac{\partial f_1}{\partial x_3} = \left( -q_{ac \max} k_{ia} k_4 (k_a + x_2 / x_3) x_2 / (k_{ia} + x_2 / x_3) x_3^2 + q_{ac \max} k_{ia} k_4 (k_a + x_2 / x_3)^2 x_2^2 / (k_{ia} + x_2 / x_3) x_3^3 \right. \\ \left. + q_{ac \max} k_{ia} k_4 (k_a + x_2 / x_3) x_2^2 / (k_{ia} + x_2 / x_3)^2 x_3^3 \right) x_1$$

$$\frac{\partial f_2}{\partial x_3} = q_{ac \max} x_2 k_{ia} (k_a + x_2 / x_3) / x_3^2 (k_{ia} + x_2 / x_3) x_1 - q_{ac \max} x_2^2 (k_a + x_2 / x_3)^3 x_1 / k_{ia} x_3^3$$

$$- q_{ac \max} k_{ia} x_2^2 (k_a + x_2 / x_3) / x_3^3 (k_{ia} + x_2 / x_3)^2 x_1$$

○ Sous-régime 2 :

Les expressions des éléments de la matrice  $\partial f / \partial x$  dans le cas du sous-régime 2 sont :

$$\frac{\partial f_1}{\partial x_1} = q_{s \max} k_1 / k_s + k_{os} / k_4 (q_{o \max} (k_{io} + x_2 / x_3) / k_{os} k_{io} - q_{s \max} / k_s) k_{oa}$$

$$\frac{\partial f_2}{\partial x_1} = k_{os} / k_{oa} (q_{s \max} / k_s - q_{o \max} (k_{io} + x_2 / x_3) / k_{os} k_{io})$$

$$\frac{\partial f_1}{\partial x_2} = -1 / k_4 q_{o \max} k_{io} x_3 / (k_{io} + x_2 / x_3)^2 k_{oa} x_1$$

$$\frac{\partial f_2}{\partial x_2} = 1 / k_{oa} q_{o \max} x_3 x_1 k_{io} / (k_{io} + x_2 / x_3)^2$$

$$\frac{\partial f_1}{\partial x_3} = 1 / k_4 q_{o \max} x_3^2 k_{io} / (k_{io} + x_2 / x_3)^2 x_2 k_{oa} x_1$$

$$\frac{\partial f_2}{\partial x_3} = -1 / k_{oa} q_{o \max} x_3^2 x_1 k_{io} / x_2 (k_{io} + x_2 / x_3)^2$$

- Régime oxydo-fermentatif :

Dans ce régime, les expressions analytiques des dérivées partielles de  $f$  sont :

$$\frac{\partial f_1}{\partial x_1} = (q_{s \max} / k_s - q_{o \max} (k_{io} + x_2 / x_3) / k_{os} k_{io}) / k_2 + q_{o \max} (k_{io} + x_2 / x_3) / k_{os} k_{io} k_1$$

$$\frac{\partial f_2}{\partial x_1} = k_3 / k_2 (q_{s \max} / k_s - q_{o \max} (k_{io} + x_2 / x_3) / k_{os} k_{io})$$

$$\frac{\partial f_1}{\partial x_2} = (q_{o \max} (k_{io} + x_2 / x_3)^2 k_2 / k_{os} k_{io} x_3 - q_{o \max} (k_{io} + x_2 / x_3)^2 / k_{os} k_{io} k_1 x_3) x_1$$

$$\frac{\partial f_2}{\partial x_2} = k_3 / k_2 q_{o \max} (k_{io} + x_2 / x_3)^2 / k_{os} k_{io} x_3 x_1$$

$$\frac{\partial f_1}{\partial x_3} = (-q_{o \max} (k_{io} + x_2 / x_3)^2 x_2 k_2 / k_{os} k_{io} x_3^2 + q_{o \max} (k_{io} + x_2 / x_3)^2 x_3^2 / k_{os} k_{io} k_1 x_2) x_1$$

$$\frac{\partial f_2}{\partial x_3} = -k_3 / k_2 q_{o \max} (k_{io} + x_2 / x_3)^2 x_2 / k_{os} k_{io} x_3^2 x_1$$

### 9.3 Annexe 3 : Expression de $S_{crit}$ en fonction de $A_{set}$

Cette annexe donne l'expression détaillée de la concentration en glucose  $S_{crit}$  en fonction de la concentration en acétate  $A_{set}$ . Rappelons que le couple  $(A_{set}, S_{crit})$  présenté au chapitre 5 vérifie la relation suivante :

$$\left. \frac{q_S}{(S_{in} - S)} \right|_{S=S_{crit}} = \left. \frac{r_A}{A} \right|_{A=A_{set}, S=S_{crit}} \quad (A3.1)$$

L'expression de  $S_{crit}$  en fonction de  $A_{set}$  est donnée à l'aide grâce à la *Symbolic Matlab*. Il existe deux solutions possibles :

$$\begin{aligned} S_{crit} = & \frac{1}{2k_3(q_{Scrit} - q_{s\max})} [k_3 S_{in}(q_{Scrit} - q_{s\max}) + q_{s\max} k_2 A - q_{Scrit} k_s k_3 \\ & + (k_3 S_{in} q_{s\max} (k_3 S_{in} q_{s\max} - 2k_3 S_{in} q_{Scrit} - 2k_3 k_s q_{Scrit} - 2q_{s\max} k_2 A + 2q_{Scrit} k_2 A) \\ & + k_3^2 k_s^2 q_{Scrit}^2 - 2q_{Scrit} k_s k_2 A k_3 q_{s\max} + q_{s\max}^2 k_2^2 A^2 + k_3^2 S_{in}^2 q_{Scrit}^2 + 2k_3^2 S_{in} q_{Scrit}^2 k_s)^{1/2}] \end{aligned} \quad (A3.2)$$

$$\begin{aligned} S_{crit} = & \frac{1}{2k_3(q_{Scrit} - q_{s\max})} [k_3 S_{in}(q_{Scrit} - q_{s\max}) + q_{s\max} k_2 A - q_{Scrit} k_s k_3 \\ & - (k_3 S_{in} q_{s\max} (k_3 S_{in} q_{s\max} - 2k_3 S_{in} q_{Scrit} - 2k_3 k_s q_{Scrit} - 2q_{s\max} k_2 A + 2q_{Scrit} k_2 A) \\ & + k_3^2 k_s^2 q_{Scrit}^2 - 2q_{Scrit} k_s k_2 A k_3 q_{s\max} + q_{s\max}^2 k_2^2 A^2 + k_3^2 S_{in}^2 q_{Scrit}^2 + 2k_3^2 S_{in} q_{Scrit}^2 k_s)^{1/2}] \end{aligned} \quad (A3.3)$$

Seule la première solution est physiquement acceptable. En effet, la solution (A.3.2) représente une faible concentration en substrat, ce qui est le cas de fonctionnement des *E. coli*. La deuxième solution est élevée et de l'ordre de  $S_{in}$ . C'est vers cette deuxième solution physiquement irréalisable que la commande GMC appliquée sur le modèle simplifié (avec quatre variables d'état) converge comme mis en évidence au chapitre 5.

## 9.4 Annexe 4 : Montage expérimental

L'implantation de la loi de commande développée dans cette thèse devrait être réalisée au laboratoire de biochimie de la Faculté Polytechnique de Mons en Belgique. Pour se familiariser avec l'environnement expérimental, plusieurs séjours à Mons ont été effectués au cours de la thèse. Cette annexe présente le montage expérimental existant à ce jour. Le bioréacteur est encore en cours d'instrumentation, de sorte qu'aucune validation expérimentale n'a pu être réalisée.

### 9.4.1 Présentation du bioréacteur

La Figure A4.1 représente une photo générale de ce montage. Les différents dispositifs sont ensuite détaillés un à un.



**Figure A4.31.** Photo du bioréacteur de culture des bactéries *E. coli* et les différents analyseurs utilisés.

La Figure A4.2 représente la photo du bioréacteur *E. coli* en cours de fonctionnement. Il est posé sur une balance et lié aux différents capteurs disponibles. Il a une capacité de 2 litres.



**Figure A4.2.** Photo du bioréacteur de culture des bactéries *E. coli*.

## 9.4.2 Dispositifs de mesures

Nous présentons ici les différentes sondes disponibles au laboratoire. Ces différents dispositifs permettent des mesures en ligne ou hors ligne.

### 9.4.2.1 Mesure de la concentration en biomasse

L'analyseur « Biomass System-FOGALE » présenté Figure A4.3 donne les mesures en ligne de la capacité et de la permittivité du milieu de culture, avec un pas d'échantillonnage de 3 minutes. Cet analyseur est équipé d'une sonde diélectrique. Il a été développé par la société FOGALE nanotech et a pour avantages la mesure des cellules viables et la surveillance en ligne de la morphologie et de la physiologie des cellules.

L'obtention de la mesure de la concentration en biomasse est donnée par le graphe de conversion de pF à g/L (conversion capacité à concentration). Ce graphe est établi par les chimistes.



**Figure A4.3.** Analyseur Biomass system-FOGALE.

### 9.4.2.2 Mesure de la concentration en glucose

L'analyseur « ProcessTRACE » Figure A4.4 nous permet d'avoir la mesure en ligne ou hors ligne de la concentration du glucose en g/L au sein du bioréacteur avec un pas de six minutes.



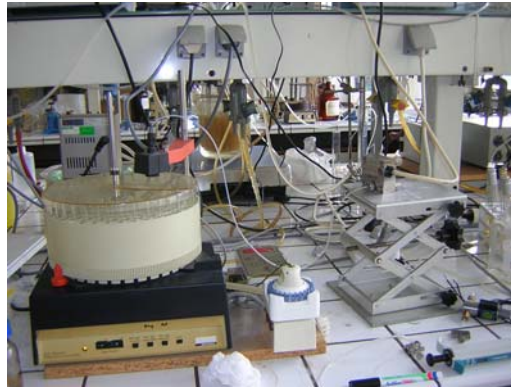
**Figure A4.4.** Analyseur 'ProcessTRACE' permettant la mesure en ligne de la concentration du glucose dans le milieu en g/L.

Cet analyseur est fourni par la société TRACE Analytics. L'analyse du glucose dans ce type de milieu de culture nécessite un prélèvement dans des conditions stériles tout en gardant les

propriétés de l'échantillon prélevé. C'est la raison pour laquelle cet analyseur dispose de trois options pour l'échantillonnage, en utilisant soit une sonde de filtration, soit une sonde de dialyse ou soit un module de déviation. Quelle que soit la sonde choisie, le prélèvement stérile est fiable. Dans notre cas, une sonde de dialyse a été choisie.

#### 9.4.2.3 Mesure de la concentration en acétate

La mesure de la concentration en acétate est réalisée hors ligne, car le projet de réaliser un capteur permettant la mesure de la concentration en ligne n'a malheureusement pas abouti à ce jour. De plus, ce capteur est très rare sur le marché. Néanmoins, nous avons mis en place un dispositif permettant le prélèvement automatique d'échantillons filtrés. Les prélèvements sont effectués chaque 30 minutes. Sachant que l'analyse de l'acétate nécessite l'absence de la biomasse dans l'échantillon prélevé, l'A-SEP<sup>8</sup> est utilisée pour une filtration tangentielle de l'échantillon qui permet de séparer la biomasse du reste du milieu prélevé. Cette biomasse est ensuite reconduite par un tuyau à l'aide d'une pompe dans le bioréacteur. Ceci évite également des erreurs sur la détermination de la valeur de la concentration en biomasse. La mesure de la concentration en acétate est ensuite réalisée hors ligne à l'aide d'un kit enzymatique de BIOSENTEC.



**Figure A4.5.** Système de prélèvement en utilisant une filtration tangentielle.

#### 9.4.2.4 Mesures et régulations du pH et de l'oxygène dissous

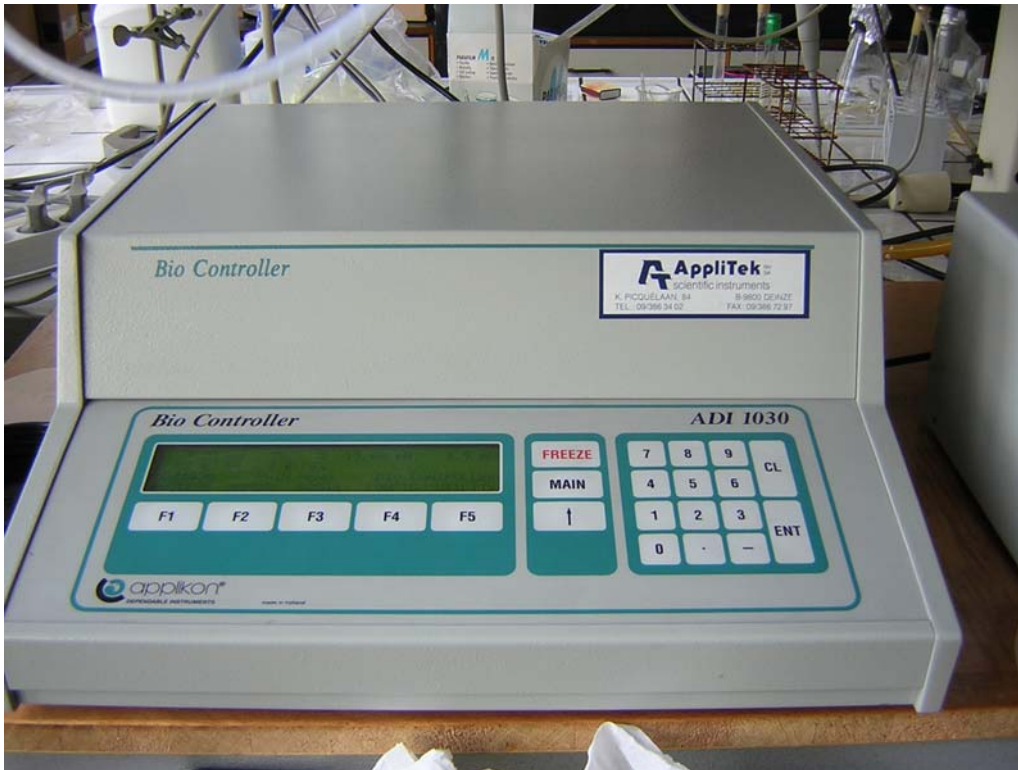
Le contrôleur « Bio Controller ADI 1030 » Figure A4.6 permet la régulation en ligne du pH et de l'oxygène dissous. La structure de commande utilisée est un PID.

#### 9.4.2.5 Mesure du poids

Le bioréacteur est disposé sur une balance (Figure A4.2), celle-ci détermine le poids du milieu de culture en l'initialisant à zéro quand le bioréacteur est complètement vide.

Toutes ces sondes sont reliées à un ordinateur où l'on peut stocker les données en ligne grâce à l'interface *Labview*.

<sup>8</sup> L'A-SEP est un dispositif permettant la filtration tangentielle. Il est commercialisé par la société Applikon en Allemagne.



**Figure A4.6.** Le contrôleur en ligne 'Bio Controller ADI 1030'.

### 9.4.3 Préparation d'une préculture

Pour pouvoir lancer une culture en batch, une étape préliminaire est indispensable pour la préparation de la préculture et la calibration des différentes sondes disponibles pour les mesures.

La préparation d'une préculture suit plusieurs étapes et dure quatre jours avant l'inoculation du bioréacteur :

#### J-4 :

- Préparation des boîtes de Pétri<sup>9</sup> avec GNO<sup>10</sup>,
- Préparation des tubes GNO,
- Autoclavage<sup>11</sup> des boîtes et des tubes afin de les stériliser.

#### Remarque :

Pour chaque batch, un tube et une boîte de pétri sont suffisants mais deux de chaque sont préparés pour des mesures de sécurité.

<sup>9</sup> Une boîte de Petri est une boîte cylindrique peu profonde, en verre ou en plastique munie d'un couvercle. Elle est utilisée en microbiologie pour la mise en culture de micro-organismes.

<sup>10</sup> La GNO est constituée de l'extrait de viande, *NaCl* et de la peptone. Dans le cas des boîtes de Pétri, la gélose est ajoutée pour solidifier le milieu.

<sup>11</sup> L'autoclavage est un cycle d'utilisation d'un autoclave. Ce dernier est un récipient à parois épaisses et à fermeture hermétique conçu pour réaliser sous pression une stérilisation à la vapeur.

**J-3 :**

- Repiquage des bactéries sur les boîtes de Pétri pour les mettre dans une étuve à 30° pendant 24h,
- Préparation des vitamines et des éléments en trace. Leurs compositions sont données aux Tableaux A4.1 et A4.2,
- Préparation du milieu de culture. Sa composition est donnée au Tableau A4.3,
- Autoclavage de 10% du volume total de ce milieu dans un erlenmeyer ensuite mise au réfrigérateur,
- Stockage au réfrigérateur des 90% restant de ce milieu.

**J-2 :**

- Prélèvement des bactéries et mise dans les tubes GNO et laisser à 30° pendant au moins 24h,
- Calibration de la sonde pH du bioréacteur,
- Autoclavage des 90% du volume total du milieu dans le bioréacteur.

Composant	Concentration (g/L)
Riboflavine(ou vitamine $B_2$ )	0,42
Acide pantothénique (ou vitamine $B_5$ )	5,4
Acide nicotinique (ou vitamine $B_3$ )	6,1
Pyridoxine (ou vitamine $B_6$ )	1,4
Biotine (ou vitamine $B_8$ )	0,06
Acide folique (ou vitamine $B_9$ )	0,042

**Tableau A4.1.** Composition de la solution des vitamines.**J-1 :**

- Calibration de la sonde d'oxygène,
- Calibration de la sonde de glucose,
- Mise à pH du milieu,
- Ensemencement de la préculture (10%) avec tubes GNO (2 erlenmeyers) et mise dans l'étuve à 30 °C avec agitation,
- Ajout des vitamines et des éléments en trace (sels minéraux) à l'aide d'une filtration stérile.

**Remarque :**

L'autoclavage des vitamines et des éléments en trace est impossible car les vitamines sont thermosensibles et les éléments en trace précipitent à chaud.

<b>Composant</b>	<b>Concentration (g/L)</b>
<i>FeCl<sub>3</sub></i>	27
<i>ZnCl<sub>2</sub></i>	2
<i>CoCl<sub>2</sub></i>	2
<i>NaMoO<sub>4</sub></i>	2
<i>CaCl<sub>2</sub></i>	1
<i>CuCl<sub>2</sub></i>	1
<i>H<sub>3</sub>BO<sub>3</sub></i>	0,5
<i>HCl</i>	100

**Tableau A4.2.** Composition des éléments en trace.

<b>Composant</b>	<b>Concentration (g/L)</b>
<i>Glucose monohydraté</i>	5,5
<i>Na<sub>2</sub>HPO<sub>4</sub></i>	6
<i>KH<sub>2</sub>PO<sub>4</sub></i>	3
<i>NH<sub>4</sub>Cl</i>	1
<i>NaCl</i>	0,5
<i>MgSO<sub>4</sub>.7(H<sub>2</sub>O)</i>	0,12
<i>Tiamine</i>	0,34
<i>CaCl<sub>2</sub>.2(H<sub>2</sub>O)</i>	0,015

**Tableau A4.3.** Composition du milieu de culture.

#### 9.4.4 Réalisation d'une expérience « culture »

La réalisation d'une culture démarre après la phase de pré-culture présentée précédemment. Le bioréacteur est tout d'abord lancé en mode batch jusqu'à consommation du glucose. Le Tableau A4.4 présente le milieu de culture au sein du bioréacteur en batch. Ensuite, la phase fed-batch est lancée en alimentant le bioréacteur par le milieu d'alimentation décrit au Tableau A4.5.

Composant	Batch (g/L)
<i>Glucose monohydraté</i>	5,5
$Na_2HPO_4$	6
$KH_2PO_4$	3
$NH_4Cl$	1
$NaCl$	0,5
$MgSO_4 \cdot 7(H_2O)$	0,12
<i>Tiamine</i>	0,34
$CaCl_2 \cdot 2(H_2O)$	0,015
<i>Eléments en trace</i>	2
<i>Vitamines</i>	2

**Tableau A4.4.** Composition du milieu de culture au sein du bioréacteur au démarrage du batch.

Composant	Concentration (g/L)
<i>Glucose monohydraté</i>	265
$NH_4Cl$	10
$MgSO_4 \cdot 7(H_2O)$	4

**Tableau A4.5.** Composition du milieu d'alimentation utilisé pour la phase fed-batch.